

What must be done about AIDS

Governments can no longer hope that AIDS will simply go away. Time could yet show that AIDS is as great a threat as nuclear war. And the time has come to act. But how?

THERE are two extreme views about the occurrence of AIDS (acquired immune deficiency syndrome) in places such as North America and Western Europe. One, exemplified by the sponsors of Proposition 64 on this week's Californian ballot, is that AIDS has arisen because some social groups follow unusual sexual practices (homosexuality) or inject themselves with drugs, that AIDS would go away if these practices were stopped and that society should meanwhile protect itself by casting out the sick. The other view is that AIDS has spread so rapidly since it was first recognized as a distinct disease in mid-1981, and is now spreading so rapidly outside the groups originally infected, that it can be only a matter of time before a substantial proportion of the general population, women and children as well as men, succumbs. If the second view should prove correct, and if no generally applicable remedy or prophylactic comes on the scene, the damage done to modern society by AIDS would compare with that expected from, say, a nuclear war between the major powers.

That is the background to the publication last week of the report of a committee sponsored jointly by the US National Academy of Sciences and the Institute of Medicine. One of the two committee chairmen is Dr David Baltimore, director of the Whitehead Institute at Massachusetts Institute of Technology and, piquantly, one of the two independent discoverers (with Dr Howard Temin of the University of Wisconsin) of the role of the enzyme reverse transcriptase in the replication of viruses whose genetic information is carried by RNA, of which the AIDS virus is one. The document (see p.3) is a model of clarity and careful assessment of daunting problems. It denies the first and simplistic view of AIDS and how to deal with it. Properly, it is agnostic on the second view of the distant future. But that does not prevent it from urging, irresistibly it would seem, that the US government should vastly increase its spending on the AIDS epidemic. In round numbers, it wants to see an extra \$1,000 million a year spent on education and public health and as much again on research.

The seriousness of the problem cannot be dodged. The US Public Health Service estimates that 9,000 US citizens will have died of AIDS during 1986, doubling the cumulative toll at the beginning of the year. During 1991, on the same set of estimates, there will be some 54,000 deaths in the United States, almost all of them recruited from the 1 to 1.5 million people thought already to have been infected. Elsewhere, the pattern is much the same, but lags behind that in the United States by a year or two. In Britain, for example, the number of known cases has doubled each year since 1983, and now amounts to just over 500, but there are some tens of thousands of infected people to fill the ranks of the dying in the years ahead.

Plague

So is this the plague returned? Much depends on what is meant by that concept, of which one essential ingredient must be that there is little that can be done about it. What should chill the spines of the US Congress elected this week is the judgement of the academies' committee that there is no quick therapeutic or prophylactic fix in sight; last month's wonder drug, AZT

(azidothymidine) is at best a palliative that that could prove unacceptably toxic, while a vaccine against the virus responsible (now called HIV) is at least five years away. Meanwhile, on the estimates of the US Public Health Service, AIDS will be responsible for between 1 and 2 per cent of deaths in the United States five years from now. The Baltimore report is right to emphasize the great social and economic cost of that toll, with its peculiar age-distribution between early adulthood and middle age, but what is the chance that the death-rate might grow to be an order of magnitude greater? That is what the US and other governments will want to know. While AIDS is so new, it is too soon to tell, but an answer will emerge over the next few years. Two technical questions are crucial.

Bad news

● What proportion of those asymptotically infected will eventually succumb? Since the introduction two years ago of blood tests for the detection of antibodies against HIV, it has become clear that infection is not immediately followed by AIDS. This is most directly shown by the retrospective examination of blood collected both in California and Central Africa in 1978, three years before AIDS was recognized for what it is, yet showing evidence of antibodies. The US Public Health Service's estimate of the development of AIDS in the United States supposes that 20–30 per cent of those with antibodies will convert to overt AIDS during the next five years, but that is but an educated guess based on five years experience of AIDS, only two of them with the benefit of antibody tests. The academies' committee prefers the estimate of 25–50 per cent of conversion to overt AIDS within 5–10 years of the appearance of antibodies in the blood. But what if the conversion rate is more like 100 per cent over, say, two decades? Nothing in the data so far collected can exclude that possibility, which would be very bad news.

● How effective is heterosexual intercourse as a means of infection? There was a sprinkling of women among the first few handfuls of AIDS cases reported in the United States, but these could at the outset be explained as exceptions to the rule that AIDS is a venereal disease of male homosexuals (or else is acquired by blood transfusion, haemophilia treatment or the use of second-hand syringes for injecting illicit drugs). But this view, never plausible, is no longer tenable. The academies' committee refers to substantiated cases in the United States of heterosexual transmission from men to women and in the other direction. In many African countries, as well as in Haiti in the Caribbean, women and men are more equally infected. In due course, if the social groups now considered to be especially at risk (male homosexuals, intravenous drug users and users of blood products) are set to one side, that will be the case in advanced societies as well. What will matter is the degree to which the infection spreads among the heterosexual population free from other complications. The academies' committee says that at present there is no worthwhile evidence bearing on this question. But if the answer should be gloomy, the implication could be that mere containment of the spread will be impractical. That, too, would be exceedingly bad news.

So how is it possible to fight an infection potentially so danger-



ous for which, now, there is neither prophylactic nor cure? The academics' committee agrees with the drafters of Proposition 64, which would make AIDS a notifiable disease, on just one point: in the short term, the only possible stratagem is to reduce the rate at which the infection can spread. (Historians will note the humiliating similarity with the stratagems used to contain the Black Death). Then two two groups part company.

The committee explicitly rejects proposals for the compulsory testing of the social groups considered most at risk (followed, perhaps, by the compulsory incarceration of those found positive in quarantine camps) not only on libertarian grounds but, more cogently, on the grounds that such hare-brained schemes could not work.

There are two grounds for this position. First, there is a variable, undetermined and perhaps long interval after infection before the appearance of antibodies (seroconversion) to HIV in a person's blood. That means that even ideally efficient tests could not identify all infectious carriers. More important, there is ample evidence that attempts at coercion in this field would be counter-productive; even now, without it, the social stigma of being infected with HIV means that some patients fail to report to their physicians and, for understandable if unforgivable motives, that physicians fail to report cases in their care to the authorities. Compulsion would be a way of making the disease seem to go away by driving it underground where it would be an even greater danger.

The academics' committee instead puts its money (or, rather, the federal government's) on education, demanding urgent efforts to enlighten not merely the groups at risk but the other members of the communities to which AIDS has or will spread. There is some good news. General enlightenment has taken the edge off the fear, rife even two years ago, that a person could catch AIDS from being in the same room as a patient. There is also some sketchy evidence that the behaviour of homosexual males has been changed by the spread of an understanding of the disease among those at risk, much of that accomplished by voluntary self-help groups. (Even so, the committee cannot decide whether the recent flattening of the rate of infection among homosexuals in New York City and California is a consequence of changed habits or, equivalently, the saturation of the infective population.) The committee is right to say that there are two other groups to which public education should be directed: intravenous drug users (who seem mostly beyond the ken of most social institutions) and those still at school on the threshold of adult sexual life. The committee might have made more of the threat posed by prostitutes.

Irony

These are daunting problems, loaded with irony. If the new US Congress takes the academics' report seriously, as it must, it will find itself having to spend as much on AIDS education as its predecessor chose to appropriate for the great war on drugs. At least it may be some comfort that, in respect of the transfer of HIV by means of used hypodermic needles, the two wars will overlap. Yet the committee is right to emphasize that, with the practice of screening blood products to eliminate blood from infected donors and the declining pace of new infection among male homosexuals, the intravenous drug users are the next most important reservoir of infection. Who will be able to reach them?

The young, even though they turn up at schools most days, are in some ways even less approachable. AIDS is a venereal disease that may be acquired in the course of sexual intercourse of any kind, but how are young people without sexual experience to be warned that it may be as dangerous as well as a natural practice? The committee comes out squarely for explicit sex education in the public schools and for the replacement of euphemisms such as 'intimate contact' by frank statements of what their users mean, in the vernacular if need be. That will mean more classroom four-letter words.

The committee has no choice, yet this is where its troubles will begin; it acknowledges that "political opposition and bureaucratic intransigence" at present limited school sex education to a few enlightened school jurisdictions, but seems not to appreciate that the opposition is as likely to harden as to melt away, now that there is AIDS as well as all those other problems. British readers of the report will know the difficulties, for this week in which their government will have placated its supporters opposed to sex education by transferring from local education authorities and head teachers to supposedly less free-thinking school governors the right to decide on sex education in the curriculum. Fair play, the politicians are in a nasty cleft stick.

The stakes are high; they risk that AIDS might spread beyond the point at which it could be contained. Their constituents, as always ready to believe that it would be simpler to get from A (now) to B (tomorrow) by starting from somewhere else, say they would prefer that teachers on the taxpayers' payroll should tell innocent children about the virtues of the nuclear family than about the condom as a barrier to venereal infection. But they, too, have no choice but to fall in with the logic of the academics' conclusion that only the most explicit sex education for the young offers hope of avoiding what may be a catastrophe. Issuing chastity belts for all children would be the other way.

From all this, it is possible that some good news may emerge. The remarkable feature of AIDS infection in the past five years is that so much has been learned about it so quickly. It is, for example, great good luck that the discovery of reverse transcriptase preceded the emergence of AIDS as a disease. It is also remarkable that so much has been done, in a mere three years, to identify and characterize the virus. The blood test, with all its defects, helps. Yet the provision of a prophylactic, let alone of a cure, will require a more sensitive and complete understanding of the human immune system than there is at present. The committee makes a strong case for giving research a chance; will governments agree?

Ignorance

Maybe, but maybe not. In Washington, the congressional committees will call the shots; they, not the likes of Baltimore, have the cheque-books in their fists. But congressmen will not be pleased, early in 1987, to be confronted with the demand of the committee academics' \$2,000 million a year for the sake of saving civilization, especially when the Strategic Defense Initiative will be asking for even more, backed by a tighter prospectus, for ostensibly the same purpose. The best way to choose between projects such as these is to calculate the downside risks. If SDI should come to nothing, the world would probably be no more dangerous than it is (and might even be safer), but if the search for treatment and prophylaxis for AIDS should fail, calamity will be on the cards.

The next few years, while there accumulate the data on which it will be possible to estimate the eventual penetration of the general population, will be crucial. It is a brief breathing space, one in which intelligent policies could help decelerate the spread of the disease and pave the way for techniques that will be needed, in any case, for decades to come. It will be a curious period for us all, not only for politicians. Somehow it will be necessary to combine a sense of urgency with a sense of the importance of the civility that makes society worthwhile. The danger is that some will be overwhelmed by the urgency of the problem to the point at which illiberal ways take hold. The latent and not-so-latent hostility to homosexuals is a constant danger. So are groundless attempts to keep children with AIDS out of the public schools. Heavy hints by insurance companies (as in Britain at the weekend) that voluntary blood tests will be taken as presumptive evidence of sexual risk, are both inequitable and counter-productive. What we must remember is that, as in the defence of Western Europe, battles that destroy what they are intended to preserve are not worth winning. □

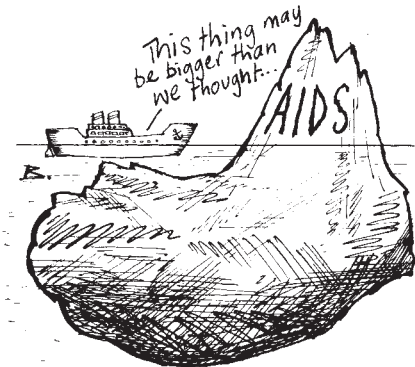
AIDS

Academies demand urgent public education

Washington

AN immediate and massive public education campaign is needed to slow the spread of acquired immune deficiency syndrome (AIDS), according to a report released last week by the Institute of Medicine (IOM) and the National Academy of Sciences*. In addition to the \$1,000 million a year required to pay for this educational campaign, the report also urges that support for AIDS research should be increased to \$1,000 million a year by 1990. Without decisive action, IOM president Samuel Thier says, the epidemic may "snowball into an even greater catastrophe".

Following its annual meeting in 1985, the Institute of Medicine determined that it should provide direction for the US effort to combat AIDS. With Thier as its new



president, the institute together with the academy made the development of recommendations for dealing with the AIDS epidemic a priority. Last week's 374-page document was produced in slightly more than seven months. David Baltimore of the Whitehead Institute and Sheldon Wolff of Tufts University Medical School co-chaired the panels responsible.

Although the report praises government support for research on the biology of the AIDS virus (HIV, or human immunodeficiency virus), the report is critical of the government's effort in public health. In particular, federal educational efforts are dubbed "woefully inadequate". Until an effective vaccine is developed, Wolff says that educating the public is the best hope for curbing the disease. The report reckons an effective campaign involving all levels of government as well as industry and private organizations will cost about \$1,000 million a year.

Almost as if anticipating the IOM report's criticism, Surgeon General C.

Everett Koop called for an increased educational effort in a report released one week before the IOM report. Declaring that "the silence must end", Koop says it is no longer possible "to sidestep frank, open discussions about sexual practices — homosexual and heterosexual". Education must be extended to prepubescent children as well as to their parents, according to the Koop report.

Wolff describes the surgeon general's report as "clearly a step in the right direction". Both reports urge the use of condoms for men who are not certain if their partners are infected with HIV.

A vexing problem in determining how best to orchestrate an educational campaign against the spread of AIDS is the lack of information about sexual practices. Baltimore suggests that a significant portion of new research money should go to social science research in that area. Sexual intercourse remains the most common means of transmitting HIV. Although the disease primarily afflicts homosexuals, the Public Health Service estimates that heterosexual transmission will account for 10 per cent of new cases by 1991.

As well as increased education, the report makes several specific recommendations about new public health efforts that should be attempted. As non-sexual transmission of AIDS occurs primarily through sharing of needles used for injection of intravenous drugs, the government should investigate the possibility of providing clean needles free of charge to users of illegal drugs. In strictly financial terms, the report says it is cheaper to treat a drug addict than an AIDS patient.

The report also calls for increased surveillance for both AIDS and AIDS-related complex (ARC) cases, and voluntary testing of serological status for individuals who may have been exposed to the virus. The report declares that mandatory blood tests are inappropriate in most cases, the armed forces being a notable exception. Calling AIDS an international problem, the IOM report urges the United States to spend \$50 million a year on efforts to control the disease throughout the world.

The report recommends that a new National AIDS Commission be established to coordinate federal efforts in combating AIDS. But Baltimore does not see this office as playing a role in directing research. He says that research proposals initiated by investigators make the best use of scientific talent in learning more about HIV, but that a directed approach

would be appropriate for a vaccine development project.

Federal funding for AIDS research has been nearly doubling in recent years. In 1984, the government spent just over \$61 million on AIDS research. For 1987, that figure will jump to \$410 million. Gary Noble, now coordinating AIDS activities for the Public Health Service, will not speculate whether that rate of increase will be sustained. Noble agrees in principle with many of the IOM report's conclusions, but says new therapeutics or unexpected discoveries about HIV could change matters.

By 1991 the Public Health Service estimates there will be 270,000 cases of AIDS in the United States. There are already an estimated 1.5 million people infected with HIV. Academy president Frank Press says there is a need for a "national strategy to respond to a national crisis". Thier hopes the reports will mark the "turning point in the nation's approach to AIDS".

Joseph Palca

AIDS drug test

SHROUDED by a veil of patent secrecy, an antiviral agent that is claimed to have more laboratory activity against the AIDS (acquired immune deficiency syndrome) virus than other anti virals is about to undergo animal toxicity tests in Britain. Porton International, the company that is developing the drug, hopes that it will pass the toxicity tests and can rapidly be tested by patients, using the US Food and Drug Administration's special fast-track procedure for potential AIDS drugs.

Porton will not, however, reveal anything of the chemical nature of the drug, which is named human immune virus antiviral. It is a natural substance extracted from a microorganism, and first emerged during a screening for antiviral compounds at the University of California (exactly where and by whom, Porton will also not reveal). Porton has acquired the rights to develop the drug which it is doing in collaboration with the Centre for Applied Microbiology and Research at Porton Down in Wiltshire, known for its expertise in the large-scale culture of microorganisms.

Peter Newmark

• THE British government, which this week set up a cabinet committee under Lord Whitelaw to look into the problems posed by AIDS is also under pressure to increase the publicity given to it and to consider a mail-drop to every home in the United Kingdom.

Critics of government strategy claim that the publicity campaign conducted in the British press since the spring has been a failure, as a high level of ignorance about the disease still prevails. Advertisements were placed in national newspapers this year, and AIDS publicity has been funded through a £2.5 million budget. □

* *Confronting AIDS: Directions for public health, health care and research.* National Academy Press, Washington, DC, 1986.

Test-ban seismology

Soviet visit to US cut short

Washington

THE unique private effort to monitor nuclear testing in the United States and the Soviet Union is off and running, with three seismic monitoring stations now established near the Soviet test facility in Semipalatinsk. But phase two, due to start next week with the arrival of five Soviet seismologists for site selection in this country, will proceed as planned. The US State Department has limited the Soviet delegation's stay to one week, and has refused to permit a planned visit to potential monitoring sites.

The monitoring effort comes from an agreement signed last spring by the Natural Resources Defense Council (NRDC), a private non-profit organization, and the Soviet Academy of Sciences (see *Nature* 321, 638; 1986). Phase one of the agreement was completed this summer when NRDC scientists established seismic monitoring stations at Karkaralinsk, Bayanul and Karasu near Semipalatinsk.

Preliminary measurements have been taken at each site, and construction is nearly completed on 300-ft boreholes for high-frequency seismometers. Returning last week from a visit to the Soviet sites, NRDC senior scientist Thomas Cochran declared the Soviets had done "an exceptional job, spending some \$150,000 for construction of each station".

The US government, by contrast, has not welcomed the NRDC effort. The administration feels that arms control issues should be carried out on a government-to-government basis, not by private organizations. The goal of the NRDC/Soviet Academy effort is "to get the verification issue out the way", says Cochran. But the State Department believes that there is a better way to monitor nuclear testing. The government's approach is a hydrodynamic yield measurement technique called CORTEX, for Continuous Reflectometry for Radius versus Time Experiment.

A coaxial cable is placed in a hole parallel to the nuclear device. After the explosion, the rate of change in the cable length is used to estimate the yield of the device. President Reagan last March offered to demonstrate CORTEX for Soviet scientists, but so far that offer has not been accepted. The State Department made viewing a CORTEX demonstration a condition for granting the two-week visa sought by the Soviet seismologists, but Cochran says that Soviet officials decided that the seismologists would be "inappropriate" for determining the value of the system.

During their visit, the Soviet delegation will also visit Teledyne Geotech and Refraction Technology in Dallas, manufac-

turers of monitoring equipment for the monitoring effort. NRDC expects to take delivery of the new equipment later this month, and after testing it should be installed on site in the Soviet Union by early next year. The Soviet seismologists will also discuss with seismologists at the Scripps Institution at the University of California, San Diego, the selection of monitoring sites near the US Nevada test facility. Three potential sites have been identified at Nelson and Railroad Valley, both in Nevada, and Deep Springs, California.

Joseph Palca

British seek funds

THE British team, invited by the Soviets to join the Americans in monitoring nuclear tests at Semipalatinsk, is attempting quickly to raise the bulk of the £200,000 it will need for the coming year.

Two committees will steer the project, one composed of eminent scientists with a reputation for raising money and the other of seismologists. Seismologists from the Universities of Leeds, Liverpool and Leicester have already offered to help.

The invitation was extended by the Soviet Academy early last month to two British scientists visiting Moscow, Dr Frank Barnaby, former director of the Stockholm International Peace Research Institute, and Dr Jeremy Leggatt, lecturer in Earth sciences at Imperial College in London.

Apart from Barnaby and Leggatt, the project team includes Dr David Davies, a seismologist and former editor of *Nature* who is now director of the Dartington North Devon Trust, Professor Martin Rees of the Institute of Astronomy, Cambridge, and Joseph Rotblat, emeritus professor of physics at St Bartholomew's Hospital Medical School in London.

The British team should be in place by mid-December, before the test-ban treaty lapses at the end of the year. If the group can attract sufficient funds, two monitoring positions will be created, manned by two people at any one time.

Because of the political aspects of the project, no research council support will be available, Leggatt says, in spite of the scientific value of the proposed collaboration between British and Soviet seismologists. It is hoped, however, that Quaker and US foundations will provide support.

The equipment will all be supplied from the West, although the Soviets will have access to data retrieved. The British monitoring positions will collect data on the crustal structure below the test sites, particularly seismic wave propagation.

Bill Johnstone

Strategic Defense Initiative

Academicians doubt efficacy

Washington

SCIENTISTS at Cornell University announced last week the results of a commissioned survey on views held by members of the National Academy of Sciences about the Strategic Defense Initiative (SDI). The principal conclusion is that 78 per cent of those responding believed that the prospects were "poor" or "extremely poor" that SDI can be made "survivable and cost effective at the margin" within the next 25 years, while only 4 per cent believed the chances of success were better than even. Ninety-eight per cent believed SDI could not provide an effective defence of the US population within 25 years in the face of Soviet attempts to thwart it, and 94 per cent believed that goal to be impossible even if the Soviets froze their arsenals at current levels.

The survey has caused some discomfort at the academy because it does not represent an official view; although the Cornell scientists have been careful to make the distinction clear, academy staff feel it is now perceived as opposed to SDI. The study, designed and conducted by the Cornell Institute for Social and Economic Research, was instigated by Peter Stein, a physics professor at Cornell, and paid for by the Ploughshares Fund of San Francisco and the MacArthur Foundation. The survey was sent to all 663 academy members in the physical, mathematical and engineering sciences, and responses were received from 71 per cent.

The study nevertheless received some influential support from academy members. A covering letter sent out with the questionnaire, signed by well-known Cornell members, urged recipients to give a few minutes to respond, saying that a "definitive poll" could provide a "useful contribution to one of the most vital science policy issues of our time".

The administration, well used to academics querying SDI, is unlikely to be immediately influenced by the survey. But Congress is another matter: it recently voted \$3,500 million for SDI research in fiscal year 1987, \$1,900 million less than the administration requested. But for a one-vote close call, the figure would have been only \$3,300 million, and doubts about SDI's feasibility played a significant part in reducing the budget. Senator William Proxmire (Democrat, Wisconsin), a prominent SDI critic who was invited to the press conference at which the survey was announced, said "the best scientific minds of the country are telling [the President] loud and clear that Star Wars won't work and that it's a waste of money".

Tim Beardsley

Robots

Britain tries to catch up

BRITAIN is to have an Advanced Robotics Research Centre to develop new techniques for automatic manufacture and to cultivate the wider acceptance of robot technology, in which Britain still lags behind many of its major industrial competitors.

The announcement of this development coincides with the publication on 5 November of a study of the uses of robots in Britain, which concludes that there are still many fewer than 1,000 manufacturing companies using robots, roughly one robot for every 40 factories. The robot population of Britain at the beginning of the year is estimated at 3,200, fewer than the annual increase of robots in West Germany alone during 1985.

The new research centre will be supported jointly by the British government and between 20 and 25 industrial partners. Possible locations include the universities of Glasgow, Salford and Edinburgh and the UK Atomic Energy Authority research establishment at Harwell. The budget will be determined by the number of industrial partners in the venture.

The plan is that the institute will have a research staff of two dozen, concentrating on advanced manipulators, sensors (visual, tactile and acoustic), navigational systems for self-propelled devices, computer systems software and standards for control and artificial intelligence or Intelligent Knowledge Based systems. The new group will also investigate advanced designs intended for use in hostile environments unsuitable for people, for example, nuclear installations.

The study published last week, which advocates the setting up of a centre along the lines now proposed, was carried out by the London-based Policy Studies Institute on the initiative of the British Robot Association with support from the Gatsby Charitable Foundation, the Department of Trade and Industry and the National Economic Development Office.

The report says that there were industrial robots in about 740 manufacturing plants at the beginning of 1986, but that most of the plants belonged to larger companies. Only 12 per cent of robot users employ fewer than 100 people. On the other hand, a third of the users are very large manufacturing plants employing more than 1,000 people.

The report says that part of the problem in Britain has been the high cost of the development of robots, the lack of technical support and the poor reliability of designs. The new robot centre is meant to offer a cure for some of these ills.

Bill Johnstone

Alaska glacier seals fate of fiord wildlife



WHEN the Landsat 5 image on the left was taken on 7 August last year, water from the Russell Fiord in Alaska (on the right in each image) could flow freely into Disenchantment Bay (left). But by 11 September this year, as shown in the image on the right, the Hubbard Glacier had sealed off the fiord, forming Russell Lake. Movement of the glacier could eventually fill Disenchantment Bay and the larger Yakutat Bay with ice.

This event is "probably the largest natu-

ral alteration in ocean, glaciers, lake and rivers in North America within our lifetimes", says US Geological Survey glaciologist Larry Mayo. The inhabitants of Yakutat on the Pacific are probably not in any danger, but saltwater fish and wildlife, including seals, trapped in Russell Lake will almost certainly die as the water becomes deoxygenated.

In these images vegetation shows up green, rock and oils a red and water/ice are in blue and white shades. □

US education

Plan to license teachers

Washington

THE US National Science Teachers Association (NSTA) is planning a certification scheme to establish basic qualifications for all science teachers. The plan has been prompted by a study showing that half of newly hired teachers in the United States are unqualified and that about a third of all science classes are staffed by unqualified teachers.

The immediate problem is that state certification requirements are now "largely ignored" in hiring and assignment decisions, according to Bill Aldridge, executive director of the association. Although participation in the certification scheme will be voluntary, Aldridge hopes that it will put pressure on states to improve and enforce their own certification standards.

The association also hopes that a further benefit of its scheme will be that specialist teachers in high schools will be more accurately assigned to suitable classes. Another survey by NSTA showed that 65 per cent of physics teachers, 52 per cent of chemistry teachers and 37 per cent of biology teachers are assigned to classes other

than those they are qualified to teach.

To qualify for NSTA certification, teachers will have to hold an appropriate bachelor's degree and will have to demonstrate a strong background in statistics, mathematics and the application of computers to science teaching. In-service training and three years experience will also be required.

NSTA's standards have been adopted by the National Council for the Accreditation of Teacher Education, and will be the basis for future decisions on the accreditation of college and university science teacher training programmes.

School administrators are likely to welcome the certification scheme, although Gary Marks of the American Association of School Administrators is concerned that the pressure to raise standards may impose greater financial burdens on school systems. He also asks whether, if certification becomes standard, school administrators will be forced to drop courses when certified teachers are not readily available.

Carol Ezzell

US oceanography

Research centre born in storm

Washington

THE US Navy's new Institute for Naval Oceanography (INO) in Bay St Louis, Mississippi, opened last month with ambitions of becoming the leading ocean modelling research centre in the 1990s. But critics are still bitter at the decision by the Navy to put the institute at the unpopular Mississippi site rather than at Monterey, California, the choice of the great majority of the experts consulted. Some say that the Mississippi site, which is far from any major university and suffers from extreme heat and humidity in the summer, will be a "major impediment" to hiring the best researchers in the field.

According to consultants who worked on the project, Bay St Louis was not even the first choice of Secretary of the Navy John Lehman. The final decision was not made until 18 months ago, and many see it as reflecting the influence of Mississippi Senator John Stennis, a powerful Democrat on the Senate Armed Services Committee. Stennis is seen as the force behind a long-term agglomeration of Navy oceanographic facilities at Bay St Louis.

Some oceanographers say that opportunities for INO collaboration with the Navy's Fleet Numerical Oceanographic Center and its postgraduate school at Monterey would have been better than with the Navy Oceanographic Research and Development Activity (NORDA), already at the Bay St Louis site. (Researchers staged a protest several years ago when part of the Naval Research Laboratory in Washington was moved to Bay St Louis to become part of NORDA.) And the Monterey climate is as famed for being pleasant as that of the Mississippi delta is for being unpleasant.

Even so, a member of Stennis's staff denies that political pressure had been brought to bear, saying the decision reflected priorities decided by the Navy several years ago.

INO will work on all aspects of ocean modelling and prediction, using supercomputers at other centres until it acquires in about two years its own 'class 7' supercomputer, the generation beyond the most advanced Cray machine, the XMP 4800.

Although the centre was formed primarily to improve predictions for naval operations, its director, Christopher Mooers, foresees many spin-offs to civilian applications. At present, ocean predictions are limited by atmospheric variability, and a major goal will be to integrate atmospheric data with oceanic models. Mooers says that working at INO will "appeal to those with a pioneering spirit".

The immediate goal, for the year or so ahead, is to provide regional predictions. The centre will be fully operational in the

early 1990s. The institute is timed to make best use of the expected welter of new remote-sensing ocean data in the next decade, when the National Aeronautics and Space Administration's TOPEX satellite will be in service and the US Navy launches its NROSS oceanographic satellites. Walter Munk, chairman of the National Academy of Sciences' ocean studies board, endorses the formation of the institute, saying of ocean modelling that "no one knows yet how well it can be done". INO will also integrate data from acoustic sensors and *in situ* measurements.

INO has been established at the National Space Technology Laboratories, a

remnant of the Apollo space programme, with an initial budget of \$4 million and a staff of 30. It is expected to expand to about 50 within five years, and will set aside about \$1 million each year to support work at other institutions. Mooers specifically welcomes proposals for collaborative research with overseas scientists.

Many academic oceanographers say quality suffers at Navy-run establishments, partly because military directors are often on short-term appointments. Mooers says that the NCAR consortium "provides our academic quality assurance". But other oceanographers are more cautious, saying the institute's potential to become a world leader will be met only if the Navy resists the temptation to keep it focused on short-term goals.

Tim Beardsley

Soviet Union

Russian criticism begins at home

THE election of Academician Gurii Marchuk as president of the Soviet Academy of Sciences has been celebrated with massive self-criticism. Instead of dwelling on recent achievements, both the report on the state of the academy presented by vice-president Vladimir Kotelnikov and the ensuing discussion stressed deficiencies in performance ranging from over-staffing with inept personnel to a shortage of computers of "average productivity".

In some cases, although no names were mentioned, the faults seem to have been at least in part associated with the attitude of the outgoing president, Anatolii Aleksandrov, towards Western science. It was Aleksandrov's contention that Soviet science can achieve world leadership in all fields on its own efforts, and that Soviet participation in international cooperation programmes should be undertaken only to further detente and to avoid duplication of effort. Now, however, Kotelnikov's report stresses that the CoCom restrictions on the transfer of Western state-of-the-art instruments constitutes a serious problem and the Soviet instrument manufacturing industry cannot satisfy the needs of the scientific community. Some time ago, said Kotelnikov, the academy raised the question of how to expand the development and manufacture of scientific instruments but it has made little progress so far, although capital investment in the academy's own instrument-making base is to increase by 150 per cent during the present five-year plan.

Similarly, Academician Boris Paton, president of the Ukrainian Academy of Sciences, urged the importance for Soviet planners of a proper understanding of technological developments in the West. Although technology is "the most holy of holies to which there is virtually no access", Soviet scientists must use all "possi-

ble and, so to say, impossible" means of establishing this information.

Paton's suggestions raise a number of interesting questions. The Soviet Union already has an extensive information network, administered, in the civil sector, by the State Committee for Science and Technology, and in the military sector by the Commission for the Defence Industry. Information on Western science and (in particular) technology, obtained by debriefing scientists returning from visits abroad, is fed into these twin networks through their permanent representatives in the research institutes. The reverse engineering of Western prototypes coming into Soviet hands is likewise carried out under the auspices of these bodies. What Paton appears to imply, therefore, is that either the academy scientists have not been sufficiently zealous in their efforts to acquire useful technological information, or that the analysis carried out by the existing information services is inadequate.

As far as Soviet science is concerned, most criticisms proved to be aspects of the endemic Soviet theme: the failure to implement research results in manufacturing and agricultural practice.

But some practical suggestions were proposed. Academician Vladimir Tuchkevich noted that, although academic research institutes are encouraged to serve the needs of industry, they have no financial incentive to do so. Instead, the research is financed by the state, and the results, including any hardware produced by the research team, either have to be handed over free of charge or sold at less than cost price. It would be better, Tuchkevich urged, if the research institute could charge industry for what it has produced, return the initial investment to the state, and use the profits for its own needs.

Vera Rich

Japanese physics

Accelerator swings into action

Tokyo

JAPAN's High Energy Physics Laboratory, KEK, has achieved a world's first by boosting the electron beam in the 3-km main storage ring of TRISTAN to 25.5 GeV during a test run of the accelerator. If tests of the positron beam go according to plan, the first collision experiments will begin before the end of the year and TRISTAN can enjoy a spell as the world's number one electron-positron collider. But plans for electron-proton collisions have been dropped in the face of competition from West Germany's HERA electron-proton collider.

TRISTAN, which has taken five years and ¥87,000 million (£388 million) to build, consists of a 2.5-GeV linear accelerator that feeds electrons and positrons into an accumulator ring where they are boosted to 6.5–8 GeV before injection into the main ring. Radio-frequency cavities in four long, straight sections of the main ring accelerate the counter-rotating bunches of electrons and positrons which can be squeezed by superconducting magnets and smashed together in the middle of each straight. But as the electrons/positrons traverse the bends of the main ring near the speed of light, energy is released as synchrotron radiation, a loss which limits the power of the accelerator.

The record just set by TRISTAN exceeds the maximum of 23.5 GeV achieved by the PETRA ring in Hamburg, although TRISTAN's electron-beam current was only 0.4 mA. The KEK scientists are racing to raise this to 2–3 mA, the minimum required for collision experiments. VENUS, the first of TRISTAN's four detectors, is "about 70 per cent complete" and the first electron-positron collisions may be performed on schedule at the end of this month, according to KEK director Tetsuji Nishikawa.

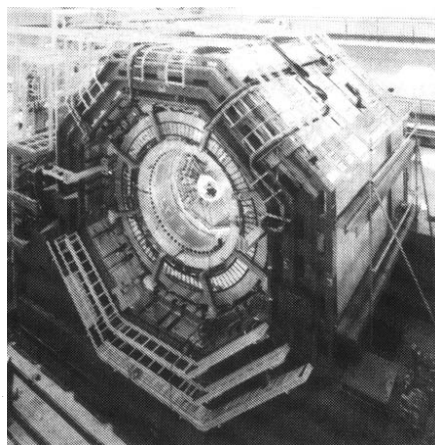
Commissioning of TRISTAN ahead of US and European rivals will be a major triumph for Japan's high-energy physicists who have been struggling to put Japan back on the map since the end of the Second World War when the country's four largest cyclotrons were dumped at sea by US marines because someone in Washington feared they could be used to make atomic bombs.

But TRISTAN's dominance may be short-lived. The Stanford Linear Collider (SLC) has already taken electrons up to 37.5 GeV towards its target of 52 GeV with the prospect of electron-positron collisions in February or March 1987.

Nishikawa, however, is confident that with its higher luminosity and wider, stable beams TRISTAN will have advantages over SLC — the beams of SLC are only about a micrometre across and there

will be difficulty in making them collide.

Among the particles TRISTAN may find are the top quark, Higgs particle, supersymmetric particles and fourth generation leptons and quarks, if they exist. But KEK scientists will also be looking for "totally unexpected and new phenomena



in the new energy frontier".

TRISTAN will continue to consume large amounts of money, mainly in the form of electricity bills. In this fiscal year (until April 1987), ¥6,000 million (£27 million) has been allotted by the Ministry of Education, Science and Culture for operation costs, enough for 1,000 hours. KEK is hoping for 1,800 hours at a cost of ¥10,000 million in fiscal year 1987.

Next March, 40 radio-frequency cavities will be added to the main ring, bringing the beam energy up to about 28 GeV, and the addition of superconducting cavities in 1988 is expected to boost this to 33 GeV, a matter of months before the LEP electron-positron collider ring at CERN is due to come on line at 50 GeV.

KEK also plans to use the main ring as a 10-GeV synchrotron radiation source, which could be another first for Japan, but the most ambitious plan of all for TRISTAN has been abandoned.

Since TRISTAN was first proposed in 1973, it had been intended to include a proton storage ring alongside the electron-positron ring to allow electrons at 30 GeV to collide with protons of 300 GeV. Space for the proton ring was made in the 3-km tunnel of the main ring and a 12-GeV proton synchrotron built at KEK in 1976 would have provided the protons. But West Germany's decision to build the larger HERA electron-positron collider at Hamburg by 1990 has apparently caused Japan to change its plans. The main reason, says Nishikawa, is that HERA, which will smash 30-GeV electrons against protons of at least 820 GeV (with even higher energies in prospect), will be much more likely to reveal quark structure.

David Swinbanks

One-way crossing

THE British electrical utility, the Central Electricity Generating Board, began last Thursday to draw some 1,500 MW of cut-price electricity from France along four pairs of high-power cables recently laid across the English Channel. The French generating system is now more than 70 per cent nuclear and has over-capacity. Under the 1–2-year deal between CEGB and its opposite number, Electricité de France, the British board will take at least 1,000 MW continuously at a cost 25 per cent below 'average' generating costs in Britain, and 'part' (typically 500 MW according to CEGB) of the remaining 1,000 MW capacity of the link at the same cost. The link is two-way, so British electrical exports to France are possible, but none is planned. The power thus to be drawn by Britain from France exceeds the capacity of a single pressurized water reactor, such as that CEGB wishes to build at Sizewell in Suffolk, and on which the planning enquiry report is now imminent.

Robert Walgate

What anniversary?

THE BBC (British Broadcasting Corporation) is this week celebrating the fiftieth anniversary of its first television broadcast with a series of past performances almost calculated to meet the Conservative (and government) party's charge that the BBC's news reporting is biased; past politicians are mocked with an even hand. But readers should judge for themselves, from this extract from *Nature* (78, 151; 1908), whether the BBC is celebrating the right anniversary. □

Distant Electric Vision.

REFERRING to Mr. Shelford Bidwell's illuminating communication on this subject published in *Nature* of June 4, may I point out that though, as stated by Mr. Bidwell, it is wildly impracticable to effect even 160,000 synchronised operations per second by ordinary mechanical means, this part of the problem of obtaining distant electric vision can probably be solved by the employment of two beams of cathode ray (one at the transmitting and one at the receiving station) synchronously deflected by the varying fields of two electromagnets placed at right angles to one another and energised by two alternating electric currents of widely different frequencies, so that the moving extremities of the two beams are caused to sweep synchronously over the whole of the required surfaces within the one-tenth of a second necessary to take advantage of visual persistence.

Indeed, so far as the receiving apparatus is concerned, the moving cathode beam has only to be arranged to impinge on a sufficiently sensitive fluorescent screen, and given suitable variations in its intensity, to obtain the desired result.

The real difficulties lie in devising an efficient transmitter which, under the influence of light and shade, shall sufficiently vary the transmitted electric current so as to produce the necessary alterations in the intensity of the cathode beam of the receiver, and further in making this transmitter sufficiently rapid in its action to respond to the 160,000 variations per second that are necessary as a minimum.

Possibly no photoelectric phenomenon at present known will provide what is required in this respect, but should something suitable be discovered, distant electric vision will, I think, come within the region of possibility.

A. A. CAMPBELL SWINSON.

66 Victoria Street, London, S.W., June 12.

European research

Applicants galore, but few grants

FOUR out of five applicants for research grants from the European Commission in Brussels fail to get a single ECU, although many of the project proposals are worthy of support, a Commission official warned British hopefuls in London last week. The occasion was organized by the Cabinet Office to give a clearer indication to British scientists and industrialists of the support available from Brussels and how to apply for it.

Apart from the shortfall between good applications and the support available, of which the Commission repeatedly complains when seeking more cash from the

Council of Ministers, there can be long delays before decisions are reached. Successful researchers often receive the cash they need only in the year after they have spent it. The Commission admits its deficiencies and has set in train a programme of 'acceleration and simplification' to tidy up its act.

According to Commission administrator Roderick Hurst, the new 'framework programme' for Commission research and development now before the Council of Ministers will also improve matters by introducing majority voting on individual research programmes, so speeding up the

time between a Commission proposal and final approval by the Council of Ministers.

Hurst says that the most frequent criticisms are of the lack of transparency in decision-making; the complexity of contract conditions (drawn up by Commission lawyers to satisfy the laws of 12 countries); the uncertain timing of calls for proposals (offering grants in specific areas); and the delays in processing and deciding to support particular proposals.

Announcements are made in the *Official Journal* of the European Communities ("which of course you all read"), but also through national Community offices, through government departments and by advertisements in journals. But one recipient of a grant says that the academic grapevine is the best means of communication.

There were a few tips for applicants last week. Nobody in paper-submerged Brussels has time to read 100-page theses, but the brief one-and-a-half page synopsis that applicants are asked to supply is critical, the sole basis for the first weeding out. Because referees may not be fluent in the language in which the synopsis is written, it must be exceptionally clear, but to the Commission it is now also essential that applications involve a collaboration among institutions from two different countries, preferably of different types (for example university and industry).

But having too many partners is impractical, listeners were advised; the usual range is three to five. Applicants, like the good examinees they no doubt often berate for the same fault, must also answer the questions on the form. "It is amazing how many people leave questions unanswered or answer them indirectly", Hurst said. Applications in non-nuclear energy are most likely to succeed (one in three accepted), followed by environment (one in four) and industrial technology (one in six). Biotechnology is at the bottom of the list, with a high demand and a low budget (only one in seven proposals in this field have been accepted).

The British Cabinet Office is preparing a booklet to assist potential applicants. According to a Cabinet Office spokesman, it is "no secret" that Britain was unhappy with the framework programme as presented to ministers earlier this year. The original documentation and legal instruments had not been a satisfactory basis for discussion, he said, though improvements are now hand. The central issue seems to be the British government's wish to apply to the Commission's proposals the disciplines it applies to its own expenditure, he said.

Robert Walgate

● The commission's presentation has been given at Leeds and Cardiff in the past week and is to be taken to Birmingham (12 November), Belfast (13 November) and Glasgow (15 November).

British research

Postdocs over the hill at thirty?

PREJUDICE in Britain against older researchers is widespread, dissuading many talented people from embarking on postdoctoral research for fear of being ruthlessly discarded within ten years. This is one conclusion from a study by the Association of Researchers in Medicine and Science (ARMS) based on a survey of applicants for recently advertised postdoctoral fellowships.

ARMS says that its survey reveals an unjustifiable prejudice in favour of people younger than 33 years. But far from attracting youth, the prejudice has caused an exodus of British talent abroad, or to other forms of research. The advertisers of the positions were generally disappointed with the number and the quality of the applicants.

The survey covered 200 postdoctoral positions, mostly in biology and medicine, advertised in the principal professional journals during three months from December 1985. ARMS says that fewer than 7 per cent of the positions were for longer than three years. The average duration of the posts on offer was 2.5 years and the average age of those appointed 28.3 years.

Most of all, the association is disturbed by the age distribution of the successful candidates. About a quarter were between 29 and 33 years of age, but only 6 per cent were older than 33. In eight cases, employers complained that many applicants were too old — in their thirties.

The association challenges this preference for younger researchers. It claims that postdoctoral scientists in their thirties are as intelligent, hardworking and competent as in their twenties, but with the advantage of greater knowledge and experience. That they should find themselves trapped in so uncertain and precarious a way of life ironically makes the whole research profession unattractive even to the younger people employers say

they wish to attract, according to ARMS.

The report of the survey includes an account of the reasons given by advertisers of the posts for the few high-quality applicants. The common theme is that better long-term treatment has attracted more researchers to industry.

One, a biochemist from Leeds, said that "anybody in molecular biology who is any good has gone abroad, into industry or has left academic circles ... it is almost impossible to fill university research posts now". Another biochemist, based in London, was astounded at the poor response. Three years ago, he said, a similar advertisement would have attracted more than 40 good applicants, but none of the 13 applicants for the post advertised this year was suitable, and it had to be re-advertised.

A Dundee physiologist endorses this view. The quality of applicants for short-term posts has fallen because career prospects have dwindled. Financial rewards are poor compared with those in industry, and better prospects await those who emigrate to the United States, Europe and Australia.

Professor Peter Campbell, deputy chairman of ARMS and professor of biochemistry at Middlesex Hospital Medical School, has been trying to persuade those who support research to change their attitudes. Campbell argues that research often depends on teams of individuals with a range of abilities and spanning many disciplines. He thinks it proper that such teams should include both people with considerable originality and those with special techniques acquired over many years. But those of the second kind tend to be discarded by the system. Campbell also blames the salary structure in British universities. Because salaries are related to age, he says, researchers often find themselves priced out of posts in their mid-thirties.

Bill Johnstone

US cancer

Can NCI cut deaths in half?

Washington

CUTTING cancer deaths in half by the year 2000 is an achievable goal, according to the National Cancer Institute (NCI). In a monograph released last week*, NCI lays out its strategy for achieving that goal. But its critics worry that the plans may do more to maintain the health of NCI than they will to reduce cancer mortality.

Edward Sondik, coeditor of the monograph, says the NCI approach has two primary focuses; reducing smoking and improving treatment. In 1985, 462,000 people in the United States died of cancer, and NCI reckons 30 per cent of those deaths are attributable to tobacco. By 2000, NCI hopes to reduce cancer mortality attributable to smoking by between 8 and 16 per cent, implying that the number of people who smoke must be reduced to a half within this decade. Although the percentage of the US population that smokes

has been declining since 1965, the rate of decline would have to double for males and considerably more than double for females for this goal to be achieved.

Reduction in mortality through treatment accounts for half the targeted decline. Using new communication tools such as the Physician Data Query (PDQ), a computerized information service, NCI hopes modern therapies will be more widely disseminated. NCI is also relying on regional cancer treatment centres and cooperating groups of physicians to test and implement new therapeutic regimens. But Sondik concedes that more than half of the projected improvements attributable to treatment must come from treatments yet to be proved effective.

In addition to smoking and therapy, the NCI report calls for changes in diet as well as more screening programmes for early detection. Reducing fat to 30 per cent of

total calorie intake and increasing fibre to 20–30 grams per day will yield an estimated 8 per cent reduction in mortality by 2000, according to the NCI report. Another 3 per cent reduction in mortality may be expected from wider use of Pap smears for all women over 20 and mammograms for women over 50.

But critics feel the NCI figures are overly optimistic. John Cairns of the Harvard School of Public Health says NCI cannot justify the expected decreased mortality projected by the report. With cancer deaths currently around 450,000 annually, Cairns says NCI projections require that more than 100,000 deaths be prevented annually through treatment. Apart from breast cancer, where improved mortality rates have been achieved, saving approximately 10,000 lives annually, Cairns wonders where the other 90,000 or more will come from.

A debate is currently raging over whether NCI is winning the war on cancer (see the *News and Views* article by Jared Diamond in *Nature* 323, 488; 1986) which was officially declared in 1971, Sondik says improved long-term survival rates show that there have been significant strides in treatment for many types of cancer. But others argue that age-adjusted mortality statistics paint a less encouraging picture and that a dramatic turnaround will be needed to achieve NCI goals.

At the heart of the debate is whether NCI is receiving too large a share of the federal research budget, and whether that share is being spent properly. Critics say it has put too great an emphasis on treatment at the expense of prevention, but Sondik counters that prevention efforts currently account for about 28 per cent of NCI outlays. Treatment programmes have been receiving approximately 30 per cent of the NCI budget.

What cannot be gainsaid is that the NCI budget is large, and growing. In 1987 NCI will receive approximately \$1,400 million, slightly less than one quarter of the total budget for the National Institutes of Health. NCI is seeking \$1,700 million for 1988. But its associate director Peter Fischinger is not sure the fight against cancer will cost much more in the future. He hopes that the mix of prevention, screening and treatment will achieve the desired results.

NCI's goal of halving cancer mortality by 2000 is not new, having been first announced in April 1984 at a congressional hearing. NCI has steadfastly maintained that dramatic reductions in cancer deaths are possible. But officials now sound a note of caution. Fischinger says the goals for 2000 are not testaments, only estimates. It will be up to Congress to decide whether those estimates are worth the hefty price tag.

Joseph Palca

*Cancer Control Objectives for the Nation: 1985–2000 NCI Monograph No. 2; 1986.

French research

Teetering on the brink

YOUNG French scientists who were offered jobs for life earlier this year by the leading French research council, the Centre National de la Recherche Scientifique (CNRS), but have since been denied this privilege, are taking their case to law.

The 522 people concerned were appointed by the usual procedures in the annual recruitment to the research council, but then were effectively dismissed by a government decree invalidating the appointing committees. The group accordingly formed a pressure group, the "Collectif des Admissibles", and appointed a lawyer to argue its case. Legal proceedings have now begun, in the highest court of the land, the Conseil d'État, against science minister Alain Devaquet, who issued the decree, and (in a lower court) against the director-general of the research council, Serge Feneuille.

Members of the group are also seeking substantial damages in private actions for loss of livelihood, expenses in relocation and other inconveniences. Meanwhile, the research council has offered one-year contracts to 285 of the 522.

The background is complex and political. The Conseil d'État ruled on 12 May that certain elections to the appointments committees of the research council had been irregular. In the event, the necessary corrections could have been made when the committees were re-elected for 1987. But the arrival of the new government and, it is said, pressure from right-wing physicians discontented with the survival of CNRS is thought to have persuaded the government to take the ruling seriously.

Devaquet's decision to dissolve the CNRS appointments committees and to annul all appointments made this year seems not to have been strictly necessary in the wake of the Conseil d'État's ruling. That is the interpretation that many French scientists are putting on the affair.

But what effect will it have on French science? A new Comité National (the body of appointments committees) is to be elected next year. Feneuille, who comes from industry, is said to want to make the stopgap short-term contracts a permanent feature of CNRS life.

Next year's Comité National will have to consider which of the 285 short-term appointees from the original 522 'admissibles' should stay; there is no guarantee that all will. Those thinking of joining the research council are also in a quandary: will the 285 have priority, affecting the numbers of new staff to be appointed?

For the longer term, French researchers are having to come to terms with lower recruitment rates than previously expected, as well as lower budgets. They will also be facing a new Comité National, which, under a new constitution, will include representatives from the whole of the university teaching community, and not just from researchers as at present. In the harder sciences, where about 60 per cent of university staff are in research and in some way connected with the research council, the effects of this change may not be too great. But in medicine, where only 30 per cent do research, the new powers in control at CNRS may be evident.

Robert Walgate

A cosmic rod in pickle

SIR—Under this title, Hoyle and Wickramasinghe¹ have tried to defend their various positions, and the claim that their earlier predictions had “wilfully and even maliciously been overlooked”.

It seems clear, however, that the real reason that their work is not widely accepted or referred to is not that they have “sinned mortally” by making a “contact between biology and astronomy”, but because so many of their various claims, put forward with so much certainty on the basis of unimpressive and even inaccurate evidence, have routinely been abandoned by themselves or proved wrong by others (see refs 2 and 3). Indeed, even before publication, their *Nature* article had been rendered obsolete by a published paper⁴ they had submitted on 29 November 1985, but did not refer to. This paper claimed hollow, spherical C_{60} (buckminsterfullerene) molecules⁵ as the explanation of the composition and ultraviolet (UV) spectrum of interstellar grains that had been earlier identified, *inter alia*, as dried hollow bacteria^{6,7}. If the absorbance in the infrared (IR) were really due to such bacteria the absorbance in the UV would be far greater than in the IR^{8,9}. They cannot have it both ways, especially as the UV and IR spectra of this C_{60} molecule have not yet been measured (R.F. Curl, personal communication).

But there are other problems. As it was part of their Fig. 1, Karim, Hoyle and Wickramasinghe had to have known that the rest of the absorption spectrum of tryptophan did not fit the curves in their Figs 3 and 5 when they claimed that “the discovery of a broad interstellar absorption feature centred on the λ 2,800Å in the extinction curve of starlight confirms the presence of proteinaceous material in grains”¹⁰. In ref. 12, they must have deliberately cut off the low wavelength part of their own (mislabelled) curve in Fig. 1 to make it fit the curves in Fig. 4 to support one of their many now abandoned theories. (For further comment, see refs 2–4, 8 and 11.) Also, Hoyle *et al.*⁶ knew that my colleagues and I¹ had never claimed the existence of a peak for *Escherichia coli* at 220 nm, and so on. This type of scientific behaviour is clearly unacceptable.

In the past few years, many comments on their publications have appeared. For example, Greenberg¹³, McLachlan and Nandy^{14,15} and Savage and Sitko¹⁶ have shown that the claimed¹⁰ absorption feature at $\lambda = 2,800\text{Å}$ was produced by saturation effects in overexposed spectra and was flagged as such in the data bank.

Thus, the long history of unwarranted claims by Hoyle and Wickramasinghe means that I, and many other scientists, will not accept anything they claim, be it

about darwinism, viruses, bacteria, *Archaeopteryx*, or even interstellar grains or steady-state cosmology, unless and until it is demonstrated by others with far better batting averages and about whose works it cannot be said “That’s not cricket”. By the way, astronomers and chemists, also, have “rods in pickle”^{*} for them too.

R.E. DAVIES

*Departments of Animal Biology, and
Astronomy and Astrophysics,
University of Pennsylvania,
Philadelphia, Pennsylvania, 19104, USA*

1. Hoyle, F. & Wickramasinghe, N.C. *Nature*, **332**, 509–511 (1986).
2. Davies, R.E., Delluva, A.M. & Koch, R.H. *Nature* **311**, 748–750 (1984).
3. Davies, R.E. *Panspermia: Unlikely, Unsupported, but Just Possible*. Preprint 36th Congress of the International Astronautical Federation, Stockholm, Sweden, October 7–12 (1985). Int. Astro. Fed. (Pergamon, Oxford).
4. Hoyle, F. & Wickramasinghe, N.C. *Astrophys. Space Sci.*, **122**, 181–184 (1986).
5. Kroto, H.W., Heath, J.R., O'Brien, S.C., Curl, R.F. & Smalley, R.E. *Nature* **318**, 162–163 (1985).
6. Hoyle, F., Wickramasinghe, N.C. & Al-Mufti, S. *Astrophys. Space Sci.*, **98**, 343–352 (1984).
7. Hoyle, F., Wickramasinghe, N.C. & Al-Mufti, S. *Astrophys. Space Sci.*, **110**, 401–404 (1985).
8. Davies, R.E., Delluva, A.M. & Koch, R.H. in *The Search for Extra-terrestrial Life — Recent Developments Int. astr. Un. symp. 112* (ed. Papagiannis, M.D.) 165–169 (Reidel, Dordrecht, 1985).
9. Yabushita, S., Wada, K., Takai, T., Inagaki, T., Young, D. & Arakawa, E.T. *Astrophys. Space Sci.*, **124**, 377–388 (1986).
10. Karim, I.M., Hoyle, F. & Wickramasinghe, N.C. *Astrophys. Space Sci.*, **94**, 223–229 (1983).
11. Koch, R.H. & Davies, R.E. *Astrophys. Space Sci.*, **100**, 425–426 (1984).
12. Karim, I.M., Hoyle, F. & Wickramasinghe, N.C. *Astrophys. Space Sci.*, **100**, 431–435 (1984).
13. Greenberg, J.M. *Observatory*, **104**, 134–135 (1984).
14. McLachlan, A. & Nandy, K. *Observatory*, **104**, 29–31 (1984).
15. McLachlan, A. & Nandy, K. *Astrophys. Space Sci.*, **109**, 399–402 (1985).
16. Savage, B.D. & Sitko, M.L. *Astrophys. Space Sci.*, **100**, 427–429 (1984).
17. Partridge, E. *A Dictionary of Slang and Unconventional English* 7th edn (Macmillan, New York, 1972).

^{*}For those not familiar with this somewhat obscure term, a “rod in pickle”, which Hoyle claims to have “not a few of”, does not mean a venerably infected penis (UK), a good horse that is being either reserved or nursed for a sure win (Australia), a pistol in a plight (USA), but a prospective punishment or scolding (UK)¹⁸.

Indian collaboration

SIR—The news item “Indian science: Foreign software giants pounce” (*Nature* **323**, 102; 1986) states “Astra’s subsidiary in Bangalore has taken the cream of Indian biochemists, including the chief biotechnologist of the Indian Institute of Science”. As I am a biochemist who has acquired the label “biotechnologist” and am intimately involved in the organization of the Astra centre, I wish to say that the statement is not correct. I am not on the payroll of Astra nor have any of my colleagues left the institute to join Astra. The core group at the centre consists of fresh PhDs who were ready to leave the country and take up postdoctoral fellowships in the United States. Besides, Astra is trying to bring back qualified Indians from abroad, so the centre is helping to prevent

a brain-drain, both internal and external.

Astra has given funds to the Indian Institute of Science, enabling me to further my basic research, which I have been carrying out for several years. At the same time, we have an opportunity to interact with the centre on problems related to the alleviation of human suffering. The Astra centre seems to us to be a unique, open and fair collaboration between an academic and an industrial research institution. The Astra centre in Bangalore is not a subsidiary of Astra Sweden. The centre will develop its own know-how and is recognized by the government as a non-profit research organization. It is registered as a society in Karnataka State. The government of India has laid down conditions to protect the Indian interests in terms of publications, patents, royalties and the conduct of research.

There is much talk about the virtues of collaboration between industry and academic research. There is also criticism that scientists in India have not been doing relevant research and that they should have discovered cures for tropical diseases a long time ago. Now that scientists are trying out a new model of collaboration and when the government has done what it can to protect Indian interests, it is a pity to lump together and to condemn all multinational efforts. I wish K.S. Jayaraman had made a deeper study of this collaboration model before offering his comments or those of his critics, which could harm not only the institution but the entire philosophy.

G. PADMANABAN

*Department of Biochemistry,
Indian Institute of Science,
Bangalore 560 012, India*

Faraway quasars

SIR—You quote me (*Nature* **323**, 193; 1986) as saying that it is only a matter of time before very long baseline interferometry yields the absolute distances to at least one quasar with an interposing gravitational lens. Alas, were that only true. What I said, in fact, was that we know of no way to determine directly the distances to quasars. But if a gravitational lens between the quasar and us forms multiple images, then, if the quasar cooperates, for example by changing its luminosity sharply, it is possible, at least in principle, to determine the differences in the times taken by light to propagate from the quasar to us along the paths for the different images. I stated my opinion that it would just be a matter of time before an imaged quasar cooperated sufficiently for us to determine such time differences.

IRWIN SHAPIRO

*Harvard-Smithsonian Center
for Astrophysics,
60 Garden Street,
Cambridge, Massachusetts 02138, USA*

Physicists about to hi-jack DNA?

Unless molecular biologists are careful and energetic, they will find that the mechanical and thermodynamic properties of DNA have been taken over by physicists.

MOLECULAR biologists, if pressed, will readily agree that more should be done to throw light on the properties of naturally occurring macromolecules in the molecular environment in which they ordinarily exercise their biological activity, usually that of water. More than that, much is being done, experimentally and otherwise. X-ray diffraction measurements of, say, crystalline proteins are not always the abstract studies they may seem, given the large water content of many specimens. The application of nuclear magnetic resonance techniques to enzyme molecules may indirectly say much about water interactions. The abiding obstacle, in these connections and others, is that the important molecules are so often complicated in their own right, even without the extra complication of their external interactions. People sigh for simpler systems.

That is where physics may have the edge. There is probably a great deal in the opinion that biologists are reconciled to having to tackle the problems with which nature confronts them, but that physicists are free to pick and choose, tackling only the problems they can reasonably hope to solve. Put starkly, like that, the distinction is obviously a simplification. (Sir Peter Medawar's *The Art of the Soluble* is, in any case, a biologist's book.) But the caricature does seem to illustrate a difference of style, which is in turn nicely illustrated by a neat analysis of the properties of DNA molecules in solution by L.L. Van Zandt of Purdue University appearing in *Physical Review Letters* (57, 2085; 1986).

The starting point for Van Zandt's analysis is an account which appeared two years ago (Edwards, G.S., Davis, C.C., Saffer, J.D. & Swicord, M.L. *Phys. Rev. Lett.* 53, 1284; 1984) of the absorption of microwaves in solutions of DNA. Part of the motivation of that work was the belief that measurements might throw light on suggestions that microwave absorption can cause biological damage even when the amounts of energy involved are very small, perhaps because of some resonance between the microwave frequency and a vibrational frequency of some vital and susceptible molecule.

But how could such a molecule, say a DNA molecule, even if characterized by well-defined frequencies of vibration, accumulate substantial amounts of energy in these vibrational modes while surrounded by water molecules whose effect would be that vibrations of any kind would

be efficiently damped? Edwards and his colleagues were sustained by their earlier observation of a dramatic increase in the gross absorption of microwaves by solutions of DNA from *Escherichia coli* in the presence of the endonuclease DNase I. They inferred that the continuing degradation of long molecules would assure a constant supply of shorter molecules capable of vibrating resonantly with external microwaves.

So why not test this possibility by using the techniques of genetic manipulation to make pieces of DNA whose size is well defined? Edwards and colleagues did just this and made two remarkable discoveries. First, well-defined DNA molecules, whether circular or linear, do indeed show resonance absorption of microwaves in the frequency range from 1 to 10 GHz. Second, the absorption peaks are surprisingly sharp, suggesting ineffective damping by the surrounding medium.

Why that should be is for Van Zandt to say, but the numbers are interesting in themselves. The measured damping implies that the relaxation time, that needed for a vibrational impulse to be dissipated, is of the order of 300 picoseconds. But the velocity of sound in DNA is within 15 per cent of 2 km s^{-1} . One consequence is that vibrational excitations can travel only about 1,500 base pairs along a linear DNA molecule before being dissipated, which in turn implies that coherent vibrations over much greater distances will not easily happen. That is why Edwards *et al.* found that their resonance frequencies were usually harmonics of a fundamental.

Van Zandt has had the benefit of an even more striking demonstration of anomalous resonance in DNA (by the Edwards group) based on measurements of DNA molecules with 5,480 base pairs derived from a circular plasmid in which supercoiling has been relieved by the use of a topoisomerase. The remarkable feature of the measurements is that the strength of the absorption (four peaks between 3 and 9 GHz) increases with increasing frequency, the opposite of that expected of an efficiently damped system.

The explanation can be made simple enough to suit any physicist. First, the vibrations responsible for the resonances are simple organ-pipe vibrations — standing compressional waves set up around the circular channels of the plasmids. A few simple properties of the solvent, water, are relevant. Energy dissipation in pure

water arises because of the stretching and distortion of hydrogen bonds, but because these form and reform in times of the order of 10^{-11} seconds, they should be measurable only at frequencies of 50 GHz and above. Van Zandt's question is how circular DNA can so profoundly modify this behaviour.

The molecular model is simple. Suppose that there is a single layer of water molecules more tightly bound to the DNA than to the bulk of the solvent, so that it will move with the molecule if that should vibrate, but at the same time be better adapted than DNA to sense and be influenced by the molecular motions in bulk liquid. Van Zandt constructs the crudest of all possible mechanical models to represent this state of affairs. Take two masses, one representing the mass of DNA (per unit length) and the other the mass of the bound water layer (again per unit length) and imagine them connected by an elastic spring. Let the second mass (the bound layer of water molecules) be damped, and calculate the absorption of the system as a function of microwave frequency.

Sceptics will say that even this simple system has enough adjustable constants to fit pretty well any measured curve. Van Zandt plays fair by using real values (the mass of DNA and the bulk viscosity of water) to fit as many of them as possible. Even the allowable frequencies of vibration are taken from measured velocities of sound in dry DNA (and the condition that there must be an integral number of wavelengths around a single circle of DNA). The result is a remarkably good fit of the position of the four absorption peaks, but the measured peaks are much broader than predicted.

Where this will lead is anybody's guess. Van Zandt promises a more detailed account of the structure of the solvent. The discrepancy between the width of the measured and predicted peaks cannot be made to go away by adjusting (and increasing) the damping constants in the model, for that has the effect of changing the relative strength of the peaks. Van Zandt thinks that even a circular plasmid whose supercoiling has been relieved may have some residual tertiary structure, but there are others who will see the difference as an opportunity to learn more about the solvent interactions. How soon will it be before molecular biology laboratories equip themselves with a klystron and some waveguide? **John Maddox**

Five-fold symmetry

Millimetre-sized quasicrystals grown from aluminium alloy

Kevin Knowles

QUASICRYSTALS, materials with 'forbidden' five-fold symmetry, were discovered two years ago, but their investigation has been dogged by the lack of large crystals. The critical X-ray and electron diffraction data which demonstrate five-fold symmetry are from very small quasicrystal particles, complicating interpretation of diffraction intensities which so far do not seem to be consistent with any clear atomic arrangement. But investigations should take a leap forward with the report

The T2 phase is not only of academic interest — it occurs as a stable grain boundary precipitate in commercial Al–Li–Cu–X alloys such as those now being developed by aluminium manufacturers for light aerospace materials with consequent savings in fuel bills. The dimensions of the Al–Li–Cu T2 grains grown by Dubost *et al.* under conditions of very slow cooling, allowing only a few nuclei to grow, contrast with the precipitates of micrometre or sub-micrometre dimensions produced from other icosahedral phases, for which the cooling rates needed to produce the phases are typically of the order of 10^7 K s^{-1} , similar to the fastest quench rates reported for obtaining crystalline phases, but far slower than quench rates for metallic glass formation. Indeed, these T2 grains are sufficiently

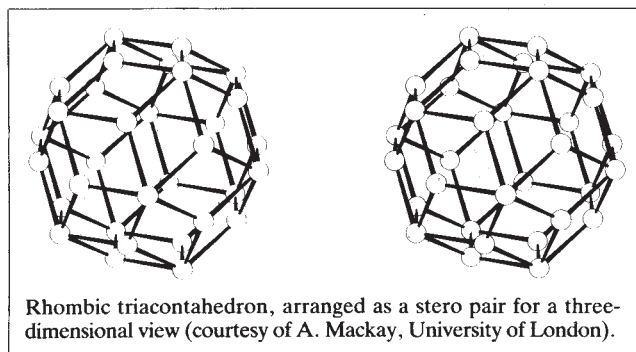
in rapidly solidified icosahedral phases. Will this also be found to be the case in the large T2 Al–Li–Cu grains where presumably equilibrium is more easily reached?

The mechanical properties of this phase are particularly relevant. Theoretical considerations by Lubensky *et al.*¹⁵ on the basis of hydrodynamic arguments predict that quasicrystals are inherently extremely brittle as dislocation motion is virtually inhibited by the slowness of phason modes in these materials.

On this basis, Lubensky *et al.* suggested that it would be virtually impossible to fabricate and manufacture macroscopic quasicrystals (such as the ones grown by Dubost *et al.*), although they recognized that the dynamics of dislocation motion in quasicrystals would alter radically if the phason modes were rendered non-hydrodynamic. It is therefore of interest to discover whether the T2 phase is indeed extremely brittle or whether it is instead merely similar to other intermetallic phases, which in general are characterized by high hardness and limited ductility. Dislocations, if they exist in icosahedral materials, must, on the basis of current electron microscope evidence, be either extremely rare or otherwise not exhibit the characteristic strainfield contrast seen in crystalline materials.

It seems clear that the report of Dubost *et al.* will stimulate more work on icosahedral materials, particularly on the thermal stability and energetics of the phase formation in Al–Li–Cu. We can only imagine where our understanding of icosahedral materials would be if more than 30 years ago Hardy and Silcock¹⁶, the original discoverers of the T2 phase, had had an electron microscope and had been able to obtain similar icosahedral electron-diffraction patterns to those found by Shechtman *et al.*² in Al–Mn. □

1. Dubost, B., Lang, J.-M., Tanaka, M., Sainfort, P. & Audier, M. *Nature* **324**, 48 (1986).
2. Shechtman, D., Blech, I., Gratias, D. & Cahn, J.W. *Phys. Rev. Lett.* **53**, 1951 (1984).
3. Mackay, A.L. & Kramer, P. *Nature* **316**, 17 (1985).
4. Mackay, A.L. *Nature* **323**, 293 (1986).
5. Horn, P.M., Malzfeldt, W., DiVincenzo, D.P., Toner, J. & Gambino R. *Phys. Rev. Lett.* **57**, 1444 (1986).
6. Stephens, P.W. & Goldman, A.I. *Phys. Rev. Lett.* **56**, 1168 (1986).
7. Socolar, J.E.S. *J. de Physique* **47**, Coll. C3 p.217 (1986).
8. Chen, C.H., Joy, D.C., Chen, H.S. & Hauser, J.J. *Phys. Rev. Lett.* **57**, 743 (1986).
9. Shechtman, D. & Blech, I.A. *Met. Trans.* **A16**, 1005 (1985).
10. Field, R.D. & Fraser, H.L. *Mat. Sci. Engng.* **68**, L17 (1984–85).
11. Urban, K. *et al. J. de Physique* **47**, Coll C3 p.465 (1986).
12. Knowles, K.M. & Stobbs, W.M. *Nature* **323**, 313 (1986).
13. Saito, Y., Takemoto, A., Tanaka, T. & Mihama, K. *Proc. 11th Int. Cong. Electron Microsc.* p.523 (1986).
14. Robertson, J.L., Misenheimer, M.E., Moss, S.C. & Bendzsky, L.A. *Acta Metallurgica* **34**, 2177 (1986).
15. Lubensky, T.C., Ramaswamy, S. & Toner, J. *Phys. Rev. B* **32**, 7444; **33**, 7715 (1986).
16. Hardy, H.K. & Silcock, J.M., *J. Inst. Metals* **84**, 423 (1955–56).



of Dubost *et al.* on page 48 of this issue¹, which describes the growing of quasicrystals approaching a millimetre across, and raises the real prospect that clear X-ray and neutron diffraction pictures of single quasicrystals will now be possible.

One of the basic principles of morphological crystallography is that the external shapes of crystals are determined by the point-group symmetry of the underlying fundamental units of the material, the unit cells, which are typically of nanometre dimensions. Thus most scientists in their schooldays will have grown crystals of alum from solution and will have found that although some will crystallize in regular octahedra characteristic of the underlying cubic symmetry, most look irregular, but nevertheless can be described in terms of simple arrangements of the unit cell.

The quasicrystals of Dubost *et al.* are grains of an Al–Li–Cu alloy showing triacontahedral morphology. They are 'quasicrystalline' in the sense that their morphology alone would imply an underlying 532 or $m\bar{3}5$ point-group symmetry, incompatible with translational symmetry inherent in the perfect crystalline state. The grains grew from the 'T2' phase of Al–Li–Cu, one of the family of icosahedral phases whose structure and properties have been of intense experimental and theoretical interest since the initial reports of these phases in alloys of Al–Mn, Al–Fe and Al–Cr^{2–4}.

large to produce X-ray and neutron diffraction patterns from single crystals. The results from such experiments are keenly awaited, as the size of quasicrystals produced to date has allowed only powder X-ray diffraction data to be obtained.

These results should help to determine the origin of the line broadening in powder X-ray diffraction patterns that can be interpreted as evidence of either frozen-in phason strain or random icosahedral packing^{5,6}. (The simplest way to see a phason in quasicrystalline materials is as a local violation in the matching rules in a Penrose tiling⁷, switching, for example, the internal vertices of rhombic triacontahedral units within the tiling.)

The ability to produce large grains of icosahedral phase will also be important for other electron-beam work, which indicates that the chemical composition of icosahedral phases grown under rapid cooling conditions is inhomogeneous on a 5–10-nm scale. Chen *et al.*⁸ have reported changes in the concentration of Mn in Al–Mn icosahedral phase from place to place, and dark-field transmission electron microscopy reveals cloudy and speckly contrast that can vary as a function of diffraction vector, beam direction and alloy formation^{9–14}.

This work implies that structural features analogous to spinodal decomposition or anti-phase domain boundaries in crystalline materials are inherent

Kevin Knowles is in the Department of Materials Science and Metallurgy, University of Cambridge, Cambridge CB2 3QZ, UK.

Molecular biology

Manipulation just off target

R. Michael Liskay

PROGRESS has been made in several laboratories towards gene targeting — recombinational interactions between an introduced DNA molecule and its homologous sequence residing in a chromosome — in mammalian cell systems¹⁻⁴. The targeting studies of Thomas and Capecchi reported on page 34 of this issue⁵ are provocative in several respects. They indicate that DNA introduced into the nucleus by microinjection can induce mutations at high frequency in the target gene in a totally unexpected manner. The observations both sound a note of caution for some applications of gene targeting and provide new methodologies that could facilitate the efficient introduction of mutations into a particular region of a gene.

High-efficiency targeting is a potentially powerful approach for manipulating the genomes of higher organisms in a manner analogous to the molecular gymnastics that are possible in yeast, including the creation or removal of mutations in any chromosomal sequence for which one possesses a molecular clone.

Thomas and Capecchi derived recipient mouse cell lines containing a mutated *neo* gene whose normal activity confers resistance to the toxic compound G418. This particular mutant form of the gene, here called the resident gene, is an 'amber' mutation that causes premature termination of translation and a resultant truncated protein. The amber mutation (a single base-pair change) is located in the initial portion of the coding region of *neo* and its presence can easily be detected by molecular techniques. Thomas and Capecchi attempted to correct the resident *neo* gene by microinjecting a different mutant form of this gene, a deletion that removes the last 52 amino acids from the protein, directly into the nucleus of the recipient cell. At a frequency of approximately 1 per 1,000 injections they recovered cells resistant to G418. Roughly half of these lines now contained, as expected, an apparently normal *neo* gene — that is, the amber mutation had been removed. Surprisingly, in the rest of the G418^r cells the amber mutation was maintained. Eight of these exceptional, corrected genes, each representing an independent occurrence, were cloned into bacteria and their nucleotide sequences near the amber mutation were determined. Each gene now harboured a compensating mutation. Three different insertion mutations were found: insertion of G, of T or of GGCT. These insertions now allow a normally silent and out-of-phase protein synthesis start signal to be

used. A functional protein is produced that is incorrect for the first few amino acids but normal for the rest. Furthermore, the mutations occurred within the first of two directly repeated sequences (GGCTAT). That each insertion results in a tandem duplication of at least a portion of the repeat sequence leads the authors to conclude that direct repeats are a prerequisite for this phenomenon.

Control experiments, including microinjection of either the same amber mutant of *neo* or an unrelated DNA molecule, did not yield any G418^r cells. From these control injections the authors conclude that the process requires both homology and at least one nucleotide difference, or mismatch, between the interacting genes. Because of this apparent requirement for a mismatch, Thomas and Capecchi invoke the formation of heteroduplex DNA between the resident gene and the incoming plasmid molecule as an intermediate and therefore refer to this newly observed process as heteroduplex-induced mutagenesis. They point out that because restoration of a functional gene was intrinsic to their experiments, many (even most) of the mutations induced by these interactions would be undetectable, making the actual frequency of mutations much higher than the 1 in 1,000 seen.

Thomas and Capecchi also analyse how the compensating mutations result in functional protein. These experiments provide a glimpse at how the ribosome can apparently reinitiate protein-chain elongation after encountering a single stop signal, provided there exists another start signal in the immediate vicinity.

The main issue, however, is the impact that these findings have on various applications of and strategies for gene targeting in mammalian cells. One might first ask why others studying targeting in mammalian cells have not observed error-prone events. The answers seem to be relatively simple. In one recent study² direct selection for targeted integration was not applied and there was no reason to examine the targeted gene at the nucleotide level. In other experiments involving direct selection in a manner similar to that of Thomas and Capecchi, the nature and/or positions of the mutation in the resident sequence were such that the detection of error-prone events would often be precluded^{1,3,4}. For example, extensive terminal or internal deletions of the resident gene would not be prone to correction by the proposed mutagenic mechanism. Note that in the bacterium *Salmonella* there exists a possibly related phenomenon

termed selfing whereby one mutation can be rescued by the very same mutant DNA⁶. Unfortunately, these events have not to my knowledge been analysed at the molecular level. On the other hand, gene conversion, a form of nonreciprocal recombination, has been shown to act with fidelity in the several cases examined in fungi⁷⁻⁹.

There are two general implications of heteroduplex-induced mutagenesis. First, the results provide a caution to those who wish to manipulate mammalian genomes by gene targeting. It is possible that unpredictable and unwanted mutations could be introduced into a gene intended to be the target for correction. For gene therapy in humans these potential side effects should be considered serious. Second, and more positive, Thomas and Capecchi's observations suggest new methods for directly mutating mammalian chromosomal sequences that differ from the approaches that are standard fare in the yeast *Saccharomyces cerevisiae*. In this yeast, specific mutations are engineered into a cloned sequence one-by-one and used to replace one-by-one the normal sequence by a targeted interaction. Based on heteroduplex-induced mutagenesis, a single appropriately constructed plasmid might prove sufficient to induce a library of mutations in a particular region of the gene. This approach might be used to advantage for mutating the control region (for example, promoter) of genes to define the location and function of *cis*-acting regulatory sequences. Such regions frequently contain repeated sequence elements and therefore should be prime targets for the proposed mutagenesis by incoming DNA.

One obstacle preventing the widespread use of targeting procedures is the propensity of mammalian cells to incorporate DNA into the genome in an essentially random fashion. The efficiency of homologous relative to nonhomologous interactions is presently being achieved at a ratio of about 1 to 100. This ratio must be improved before the mammalian genome becomes as amenable to genetic tinkering as its smaller cousin, the yeast. □

1. Smith, A. J. H. & Berg, P. *Cold Spring Harb. Symp. quant. Biol.* **49**, 171-181 (1984).
2. Smithies, O. et al. *Nature* **317**, 230-234 (1985).
3. Lin, F. L., Spearle, K. & Sternberg, N. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1391-1395 (1985).
4. Thomas, K. R., Folger, K. R. & Capecchi, M. R. *Cell* **44**, 419-428 (1986).
5. Thomas, K. R. & Capecchi, M. R. *Nature* **324**, 34-38 (1986).
6. Demerec, M. *Proc. natn. Acad. Sci. U.S.A.* **48**, 1696-1704 (1962).
7. Roman, H. *Cold Spring Harbor Symp. quant. Biol.* **21**, 175-185 (1956).
8. Case, M. & Giles, N. *Genetics* **49**, 529-540 (1964).
9. Fogel, S. & Mortimer, R. *Molec. gen. Genet.* **109**, 177-185 (1970).

R. Michael Liskay is in the Departments of Therapeutic Radiology and Human Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510-8040, USA.

Elementary particles

Colour theory in a spin

Harry J. Lipkin

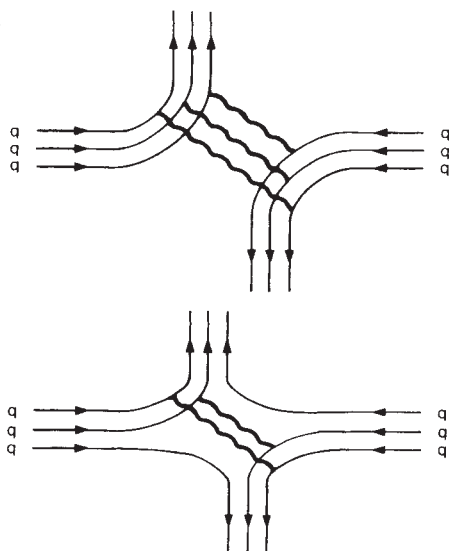
EXPERIMENTS in which beams of high-energy polarized particles are scattered by polarized protons¹⁻⁵ continue to puzzle, confuse and cause controversies among high-energy physicists. New experiments show a strong dependence of the scattering processes on the directions of the proton spins, which is not simply explained by accepted theoretical models based on quantum chromodynamics (QCD). The original establishment view was that little could be learned from polarization experiments because theory predicted no spin dependence at high energies. Subsequent investigations with QCD predicted some spin dependence, but not enough to explain the experiments. Thus some experimenters claim that they violate the predictions of QCD. The most recent experiment¹ suggests that this strong dependence suddenly disappears mysteriously at a higher energy. This causes even more confusion.

Spin physics was originally considered trivial at high energies, because of helicity conservation^{6,7}. Particles with spins, such as electrons, neutrons, protons and quarks, are called right-handed or left-handed, respectively, if their spins are oriented parallel or antiparallel to their direction of motion. QCD predicts that this handedness (called helicity) is conserved at high energy; that is, it does not change in collisions at high energy, although it may do so at low. Thus, QCD seems to predict that protons emerge from any collision with their original helicity, and that no further useful information can be obtained from any spin measurements.

The argument that helicity conservation makes spin physics trivial is strictly valid only for elementary spinning particles. In the 1950s physicists believed that protons and neutrons were elementary particles and that Yukawa's meson theory of nuclear forces explained nuclear structure and the forces holding nuclei together, just as quantum electrodynamics (QED) explained atomic structure and the forces binding atoms together. Soon many new particles, now called hadrons, were discovered and it became clear that none was elementary. We now believe hadrons are built from smaller building blocks called quarks, bound together by QCD forces. The proton is no longer elementary but contains several quarks spinning in opposite directions. Quarks with opposite helicities can be exchanged between the two protons, thus changing the proton helicities even if the helicity of each quark is unchanged. Spin measurements therefore enable us to explore the

internal helicity structure of the proton.

Spin effects played a crucial role at a very early stage⁶⁻⁸ in providing experimental tests of the quark structure of hadrons. Already in 1945 the anomalous magnetic moment of the proton was interpreted as evidence that the proton was not an elementary particle like the electron, satisfying the Dirac equation. The quark model gave the first satisfactory explanation of the magnetic moments of the proton and neutron, by having these particles made from two kinds or 'flavours' of quarks, now called up and down, with electric charges $+2/3$ and $-1/3$ of the proton charge, respectively. The proton is



Two protons, each consisting of three quarks (q) collide at high energy and are scattered through a right angle. Top, direct scattering: three high-energy gluon exchanges (wavy lines) between quarks in the two protons cause the deflection. Bottom, exchange scattering: in each proton, internal motions have already produced a quark moving in the final direction, and only two high-energy gluon exchanges are required. The amplitudes of these diagrams are added to calculate the probability of the whole process.

made of two up quarks and one down quark; the neutron of two downs and one up, thus giving the observed charges of $+1$ and 0 . Because all quarks have the same spin as the proton, the right total spin for the proton and neutron is obtained by having the up and down quarks spin in opposite directions, giving a total spin which is the difference between the up and down quark spins. But quarks with opposite electric charges spinning in opposite directions have their magnetic moments all pointing in the same direction and add up to a large value. This explains why the

proton with a positive charge has a magnetic moment roughly three times the Dirac value and the neutron with no charge has a magnetic moment corresponding to roughly double the Dirac value for a particle with negative charge. A quantitative calculation of the ratio of the proton and neutron moments with this picture gives the value $-3/2$, in remarkable agreement with experimental results.

The first excited states of the neutron and proton, now called Δ , had a spin three times the spin of the proton. These are now described as being made of the same three quarks with all spins parallel. There were also new companions of these particles, originally called strange particles. The greek letters Λ , Σ and Ξ denote the different strange particles with the same spin as the proton; an asterisk is added to denote the excited states. These are now believed to contain one or more quarks of a third flavour called strange quarks. In 1966 Federman, Rubinstein and Talmi⁹, and independently Sakharov and Zeldovich^{10,11}, calculated the excitation energies of the parallel-spin states using an extremely simple model of the spin-dependent forces based on the analogy with hyperfine interactions in atomic physics. These authors obtained relations between the masses which were in excellent agreement with experiment, for example

$$\Delta - \text{proton} = (1/2)(2\Sigma^* + \Sigma - 3\Lambda)$$

The experimental values of the two sides of this equation are 307 and 294 MeV, thus supporting the assumption that the spin forces between quarks are like the hyperfine interactions in atoms.

There is now an entire family of hadrons called baryons, all, like the proton and neutron, made from three quarks. There is also another family of hadrons called mesons, made of a quark and its antiparticle the antiquark. These include the pion (π) which has the two spins antiparallel, its strange partner the kaon (K) and their lowest excited states, denoted by ρ , ω and K^* , with their spins parallel. Sakharov and Zeldovich^{10,11} used their assumed hyperfine interactions to obtain a remarkable relation between meson and baryon masses

$$\Lambda - \text{proton} = (3/4)(K^* - \rho) + (1/4)(K - \pi)$$

The experimental values of the two sides are 177 and 180 MeV, giving striking support to the assumption that mesons and baryons are made of the same quarks.

There are many more examples of how spin effects have helped to reveal crucial clues to hadron structure which eventually led to QCD. I will now consider the attempts to understand spin effects in proton-proton scattering at high energies.

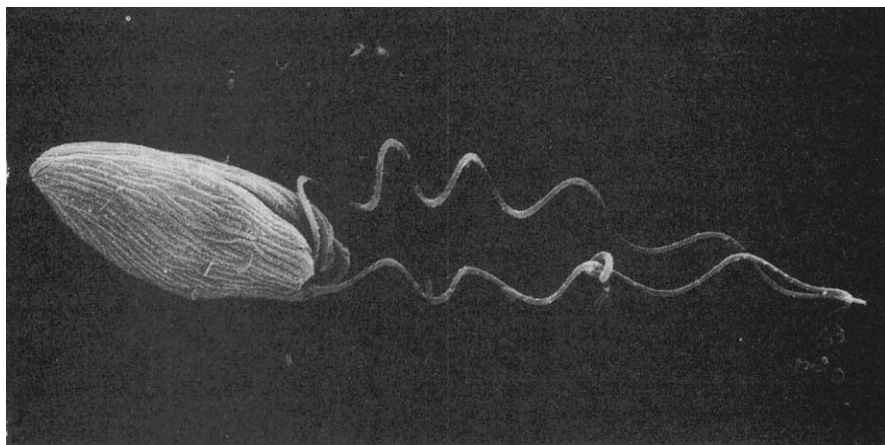
Perhaps some future article will begin: "Back in the 1980s physicists believed that quarks were elementary particles and that QCD explained hadron structure and the

forces binding quarks together". These speculations that even quarks and leptons may be composite systems can be avoided by assuming that all hadron structure physics is described by QCD just as all condensed matter physics is described by QED. But no one knows how to calculate the properties of a type II superconductor from first principles with QED; similarly there is no clear recipe for the application of QCD to the description of hadron structure and the prediction of experimental results. In both cases the systems and the dynamical equations are too complicated. Theorists use QCD hand-waving arguments to pick a principal or 'leading' contribution for calculating experimental predictions. If a prediction agrees with experiment, they are happy and write a paper. If it disagrees with experiment, they write a paper anyway, and blame the disagreement on nonleading effects. If there are no experimental results available, they publish their result as a QCD prediction. If subsequent experiments disagree with this prediction, experimenters often claim that their results disagree with QCD, rather than with a particular choice of a leading term.

Experiments in the 1970s which scattered high-energy electrons from protons showed that at high momentum transfers the proton behaves like an assembly of free quarks with very weak interactions, even though the interactions between quarks are very strong at low energies. The fact that these interactions become much weaker at high energies has been given the name asymptotic freedom and shown to be predicted by QCD. QCD arguments suggest that experiments at sufficiently high energies can be treated with the use of the leading terms in an expansion called perturbative QCD, because the interaction has become weak enough to be treated by the same perturbation techniques used in QED. But no one has a convincing prescription for where high enough energies begin, and there is a running controversy on this point between different groups of theorists.

The recent polarization experiments are not explained by treatments using any of the obvious leading terms. The first experiments showing nontrivial spin effects motivated new calculations of polarization effects based on perturbative QCD suggesting that these experiments can give useful information after all about hadron structure and QCD dynamics. The simplest leading term predicted that the probability for proton-proton scattering by 90° should be twice as large if the spins of the protons in the beam and in the target were parallel than if they were antiparallel^{12,13}.

Theorists were delighted when new experiments approached just this factor of two³ and seemed to level off there, but further experiments at higher energies showed a continued increase to factors as



This unusual microorganism was recently found by Mr Harry Burton and Dr Ian Bayly in the course of an Australian Antarctic Division examination of a meromictic lake. These lakes have two well-defined liquid layers below the ice overlay; the top layer has a high oxygen content and low salinity, the lower layer has no oxygen but does have high concentrations of ammonia and hydrogen sulphide. The boundary between the layers is called the chemocline. The microorganism — currently known as Eric — is found only in the 30 cm of water directly below the chemocline. To go with this habitat, it has an unusual lifestyle; its outer surface is covered with bacteria, which appear to cling together with 'wing-like' structures. The bacteria have not been found separately. The little that is known about Eric is reported in *Science News, Fiji*; the Australian team are at present trying to breed from him.

Rebecca Ward

large as four⁴. New models arose, all paying lip-service to QCD, but with different leading terms. The latest result¹ further confuses the theorists by giving the same value for the scattering of protons with parallel and antiparallel spins, implying a sudden disappearance of the large difference which had risen to a factor of four at somewhat lower energies. There are also other experiments showing further disagreements with theory. Single-helicity-flip transitions were observed in 28-GeV proton-proton scattering at high transverse momentum⁵, whereas the quark-exchange mechanism with helicity conservation at the quark level cannot produce a single helicity flip. An additional unexpected polarization effect was seen when ρ mesons were produced in collisions between pions and protons¹⁴. Strange particles produced in proton-proton collisions have long been known to be strongly polarized for reasons still not understood.

All these experiments suggest that there are interesting physics in these spin effects that are not understood in the simple leading terms and which may give clues to the underlying hadron structure and QCD dynamics. It is still an open question whether further experiments and analysis can reveal the relevant nonleading effects and lead to a better understanding. At present the only viable approach seems to be to continue the search for new clues in the data and try to find signals in the noise.

The basic physics and the essential complications of applying QCD to high-energy scattering can be seen by considering a colliding beam experiment in which two proton beams collide, one coming

from the left and one coming from the right, and two proton detectors look for protons scattered at an angle of 90° with respect to the incident beams, one scattered upward, say, and one scattered downward. Because each proton consists of three quarks, the collision process begins with three quarks coming in from the left and three from the right, and it ends with three quarks going upward and three going downward. A theory based on QCD must explain how these six quarks get the appropriate kicks to turn their directions of motion by 90° .

There are several possible scenarios (see figure). The direct scattering scenario has all three quarks that came in from the left going out upward, and all three quarks that came in from the right going downward, or vice versa. The exchange scenario has two of the quarks that came in from the left going out upward and one going downward, and two of the quarks that came in from the right going downward and one going upward or vice versa. In the direct scenario, each of the three quarks coming from the left must have its momentum turned by 90° by exchanging one or more gluons with quarks coming in from the right. The exchange scenario makes use of the complicated internal motion of the quarks inside each proton. At the instant before the collision, there is a certain probability that this internal motion inside the proton coming in from the left will already have two of the three quarks moving upward and one moving downward, and that the proton coming in from the right will already have two of the three quarks moving downward and one moving upward. In this case it will be

natural for the three quarks already moving upward before the collision to combine to make a proton going upward, and similarly for those moving downward.

Perturbative QCD suggests that as the interaction is weak at high momentum transfers, the scenario with the minimum number of high momentum kicks should win. The direct scenario requires high momentum kicks to turn each of the six quarks by 90°; this is very difficult at high energies. In the exchange scenario, the exchanged quarks are already going in nearly the right direction before the collision, the kick required is smaller and the QCD interaction stronger, so it is much easier to turn them around. However, the probability that they are initially moving in nearly the right direction is small. Thus two small factors enter into any theoretical comparison, the small probability of considerably changing the momentum of a quark during the collision and the probability that the quarks may initially have a transverse component of momentum in the proper direction. The QCD theory should be able to calculate both of these small factors. They will depend on the energies of the protons and on the scattering angles, which need not be 90°. The recent polarization experiments indicate that leading-term predictions in perturbative QCD disagree with experiment.

Some theorists attempting to find new non-leading effects have investigated how in the exchange scenario the quarks in a proton acquire the large transverse component of momentum necessary to be going in the right direction before the collision. The leading term in the description of the structure of the proton has all three quarks moving in the same direction. There is a tiny probability that two quarks within the proton can have a single 'hard' collision with a high momentum transfer and come apart with high transverse momenta in opposite directions. But they can also have many 'soft' collisions with smaller momentum transfers which have a much higher probability and build up to high momenta from many successive small momentum transfers. The single hard scattering can be treated by perturbative

QCD and conserves helicity. But many soft scatterings cannot be treated by perturbative QCD. Such 'non-perturbative' processes must be treated by non-leading terms and do not necessarily conserve helicity. Some model based on QCD is needed to estimate the importance of these effects, which might explain the single-helicity-flip transitions and other disagreements with perturbative QCD.

Thus, spin effects in high-energy hadron scattering are not trivial. They may shed light on how quarks are bound together by the forces of QCD to make hadrons. The problem of the description of spin effects in hadron scattering by QCD remains a challenge. □

Harry J Lipkin is at The Weizmann Institute of Science, Rehovot, Israel.

Immunology

Unravelling the control of B cells

Gerry G. B. Klaus

A CENTRAL feature of the humoral immune response is that when quiescent B lymphocytes encounter antigen they become activated, proliferate and differentiate into high-rate antibody-secreting cells. This process requires the participation of helper T lymphocytes and accessory cells such as macrophages. It has become increasingly clear that the helper function of T cells can be explained by their production of growth and differentiation factors that act on B cells. The precise functions of these molecules have, however, remained elusive until recently. Now, the complementary DNA encoding several of these lymphokines has been cloned and the sequences of two more are described in this issue^{1,2}, so the stage is set to unravel this complex process.

The existence of such factors was deduced about 15 years ago with the demonstration that culture supernatants from antigen or mitogen-stimulated mouse T cells contain factors (T-cell replacing factor(s) — TRF) which enable T-cell-depleted spleen cells to produce specific antibody responses³. Subsequently, such conditioned media revealed a bewildering variety of activities, such as activation of B cells from the resting state (B-cell stimulating factors — BSF); promotion of B-cell proliferation (BCGF); and/or differentiation to antibody secretion (BCDF). The development of monoclonal antibodies to individual lymphokines and of cloned T cells which secrete a limited array of factors has finally led to some progress in sorting out these activities. Within the past year, progress has markedly accelerated with the application of recombinant DNA technology.

This era was heralded by the cloning of the complementary DNA for murine BSF-1 (refs 4,5), followed quickly by that for the human homologue⁶. Murine BSF-1 has a relative molecular mass (M_r) of 20,000 (20K) and shows multiple activities; it acts on resting B cells and induces them to enter a transitional activation state characterized by the expression of high levels of antigens encoded by the class II major histocompatibility complex. In

combination with antibodies to antigen receptors it stimulates these cells to proliferate. BSF-1 can also act as a rather selective B-cell differentiation factor by promoting the secretion of IgG1 and IgE antibodies by B cells pre-activated with lipopolysaccharide.

A second factor (BCGF-II, $M_r \sim 45K$) was initially recognized as a growth factor for a murine B-cell lymphoma (BCL₁)⁷. It has recently been shown to cause both DNA synthesis and antibody secretion (IgM and some IgG) in pre-activated normal mouse B cells⁸. Kinashi *et al.*¹ in this issue describe the cloning of the complementary DNA encoding murine BCGF-II, and show that the product also has TRF activity.

A third molecule is the M_r 20K BCDF, which induces antibody secretion in both Epstein-Barr virus-transformed human B cells and pre-activated normal cells, but has no BCGF activity⁹. The complementary DNA for this molecule has also now been cloned by Hirano *et al.*² but the murine homologue of this molecule appears not yet to have been identified. Interestingly, this human BCDF is the only one of these three molecules that so far seems to act only on B cells⁹.

In contrast, murine BSF-1 is a growth factor for mast cells, and also for T cells, in a way distinct from interleukin-2 (ref. 10). The human homologue also has T-cell growth factor activity⁴. Murine BCGF-II also acts as an eosinophil differentiation factor (EDF, ref. 11), as does highly purified TRF (C. Sanderson, A. O'Garra and K. Takatsu, unpublished data), so there is little doubt that these three activities are produced by the same molecule. Lack of lineage specificity is common among cytokines. Interleukin-1 and interleukin-3 show diverse biological activities and interleukin-2 promotes B-cell proliferation and/or differentiation to antibody secretion (see ref. 12 for a review).

Not surprisingly, this rapid expansion of knowledge in B-cell factors has caused problems over nomenclature. BSF-1 and BCGF-II deserve to be added to the list of the three interleukins because of their

1. Court, G.R. *et al.* *Phys. Rev. Lett.* **57**, 507 (1986).
2. Gourlay, S.A. *et al.* *Phys. Rev. Lett.* **56**, 2244 (1986).
3. Crabb, D.G. *et al.* *Phys. Rev. Lett.* **41**, 1257 (1978).
4. Crossbie, E.A. *et al.* *Phys. Rev. D* **23**, 600 (1981).
5. Raymond, R.S. *et al.* in *Intersections Between Particle and Nuclear Physics* (ed. Mischke, R.E.) 1123 (AIP, Steamboat Springs, 1984).
6. Brodsky, S.J. & Lepage, G.P. *Phys. Rev. D* **23**, 2848 (1981).
7. Lipkin, J. in *School on Intermediate and High Energy Physics*, (ed. Falk, W.R.) (TRIUMF report, 1985).
8. Lipkin, H.J. *Ann. N.Y. Acad. Sci.* **452**, 79 (1985).
9. Federman, P., Rubinstein, H.R. & Talmi, I. *Phys. Lett.* **22**, 203 (1986).
10. Zeldovich, Ya. B. & Sakharov, A.D. *Yad. Fiz.* **4**, (1986).
11. Zeldovich, Ya. B. & Sakharov, A.D. *Sov. J. nucl. Phys.* **4**, 283 (1967).
12. Brodsky, S., Carlson, C. & Lipkin, H.J. *Phys. Rev. D* **20**, 2278 (1979).
13. Farra, G., Gottlieb, S., Sivers, D. & Thomas, G.H. *Phys. Rev. D* **20**, 202 (1979).
14. Barton, D.S. *et al.* *Brookhaven natn. Lab. Rep. No. BNL-34771* (1984).

multiple biological activities, and both were contenders for interleukin-4 (refs 4, 5, 11). As complementary DNA for BSF-1 was cloned first, this factor will retain that designation, and BCGF-II/TRF/EDF will presumably become interleukin-5 (as proposed by Kinashi *et al.*¹). These designations, however, have yet to be officially ratified.

Does our current knowledge about these factors produce a coherent scheme of growth and differentiation control in antigen-stimulated B cells? The answer, unfortunately, is not yet. In murine B lymphocytes it is an attractive idea that resting cells become activated by the synergistic action of antigen plus BSF-1 produced by the cooperating T cell, and then become responsive to BCGF-II later in their cycle. Perhaps the balance between these two factors, in conjunction with unknown microenvironmental influences, will then determine which classes of antibody are secreted by the clonal progeny of a particular B cell, or if these progeny will instead revert to the resting state of memory B cells. It is significant that both BSF-1 and BCGF-II are growth and differentiation factors, suggesting that clonal expansion and maturation are coordinately regulated.

Unfortunately, what we know of human B cells does not fit this scheme. First, the genetic homologue of murine BSF-1 has not yet been shown to activate resting human B cells, but rather stimulates proliferation of pre-activated cells⁶. Second, it is likely that there is yet another type of B-cell growth factor for human cells (distinct from interleukin-2). Finally, the factor described by Kishimoto *et al.*⁹ has no growth-promoting activity, suggesting compartmentalization of growth and differentiation in human B cells.

In conclusion, there still may be too many factors in the arena for comfort. Nevertheless, given the availability of the recombinant molecules, various questions can now readily be answered. For instance, are there functional B-cell subpopulations responsive to different factors? Do mixtures of individual lymphokines produce synergistic effects? The answers should not now take long to emerge. □

- Kinashi, T. *et al.* *Nature* **324**, 70 (1986).
- Hirano, T. *et al.* *Nature* **324**, 73 (1986).
- Dutton, R. W. *et al.* *Proc. Immun.* **1**, 355 (1971).
- Noma, Y. *et al.* *Nature* **319**, 640 (1986).
- Lee, F. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 2061 (1986).
- Yokota, T. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 5894 (1986).
- Swain, S.L. & Dutton, R.W. *J. exp. Med.* **156**, 1821 (1982).
- O'Garra, A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 5228 (1986).
- Hirano, T. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 5490 (1984).
- Mosmann, T.R. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 5654 (1984).
- Sanderson, C.J., O'Garra, A., Warren, D.J. & Klaus, G.G.B. *Proc. natn. Acad. Sci. U.S.A.* **83**, 437 (1986).
- Meidema, F. & Melief, C.J.M. *Immun. Today* **6**, 258 (1985).

Gerry G.B. Klaus is in the Division of Immunology, National Institute for Medical Research, London NW7 1AA, UK.

Neurobiology

Questions of taste and smell

Gordon M. Shepherd, Thomas V. Getchell and Charlotte M. Mistretta

RECEPTOR mechanisms in olfaction and taste have been tough nuts to crack for traditional methods of biochemistry and electrophysiology. The advent of molecular neurobiology and patch-clamp recordings is therefore most welcome, as was evident at a recent meeting*.

In olfaction, biochemical and molecular biological studies were made much easier by the development of isolated olfactory ciliary preparations containing enriched membrane fractions. Several molecular components localized in the fractions are believed to be associated with sensory transduction. The proposed sequence is initiated by binding of odour molecules (odogens) to membrane glycoproteins, such as gp95 (Chen, Z. & Lancet, D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1859; 1985), located in the cilia. The odogen-receptor complex is believed to activate the stimulatory GTP-binding protein (G_s) (D. Lancet, The Weizmann Institute). Although they appear to be less abundant, G_i and G_o binding proteins have also been identified in olfactory cilia using monospecific antisera (R.R.H. Anholt, Johns Hopkins University). Their functions may be related to receptor 'inhibition', or to neuromodulatory effects of compounds such as substance P (J.F. Bouvet, Université Claude Bernard) or luteinizing hormone-releasing hormone (C.R. Wirsig and T.V.G., Wayne State University). G_i -deficient pseudohypoparathyroid patients have impaired olfactory perception compared with individuals with normal G_i activity (H.N. Wright, State University of New York).

Cyclic AMP is believed to be a second messenger that is activated during olfactory transduction and leads to excitation of receptor neurones. The activity of cyclic AMP in the olfactory cilia preparations is 15 times that in the brain, and is increased substantially by GTP, its analogue GTP γ S and forskolin, which indicates that the cyclase is activated by a G protein (P.B. Sklar, Johns Hopkins University). Certain molecules evoking different odour sensations also markedly increase cyclase activity. Others, which evoke putrid or solvent odours, do not stimulate this enzymatic activity. This observation suggests that parallel, non-adenylate cyclase-dependent transduction mechanisms are operable. For example, isolated olfactory ciliary preparations also exhibit increased phosphodiesterase activity when stimulated with L-alanine

and GTP (R.C. Bruch and T. Huque, Monell Chemical Senses Center); this suggests that certain odogens stimulate phosphoinositide turnover by a G-protein-dependent mechanism.

A major question is to determine if cyclic AMP directly activates membrane channel proteins associated with the ion conductance changes responsible for the receptor potential, or if the effect is mediated by phosphorylation of specific proteins by phosphokinases in the cilia. Protein kinase C has been identified by phorbol ester binding and monospecific antisera (R.H. Anholt), suggesting that phosphorylation of the ion channels may be associated directly with conductance changes.

Patch-clamp recordings demonstrate that olfactory receptor neurones have interesting passive membrane properties that include high input resistances (2–6 G Ω) and low input capacitances 2.6–5 pF (S. Firestein and F. Werblin, University of California, Berkeley; N. Suzuki, Hokkaido University). Membrane depolarizations are associated with a rapid inward Na⁺ current followed by several outward K⁺ currents (V.E. Dionne, University of California, San Diego; S. Firestein and F. Werblin). Near-threshold concentrations of odour molecules delivered by pressure injection elicit a 6-pA generator current that depolarizes the cell by 15 mV. There is a systematic relationship between the concentration of the odogen, the magnitude of the generator current and the number of action potentials initiated (S. Firestein and F. Werblin). These results support previous evidence that there is relatively close coupling between the conductance sites associated with ligand binding and those associated with impulse initiation (L. Masukawa, *J. Neurosci.* **5**, 128; 1985).

The relationship of the intermediate molecular stages of transduction and ion conductance changes has been investigated using G protein and cAMP analogues (D. Trotier and P. MacLeod, Laboratoire de Neurobiologie Sensorielle), which enhance conductive charges, supporting the biochemical evidence for an enzymatic cascade leading to the generator current. Single channel conductances in functionally reconstituted membrane proteins isolated from olfactory cilia incorporated in bilayer lipid membranes are enhanced by odour molecules and cyclic AMP (V. Vodyanoy and I. Vodyanoy, University of California, Irvine).

A protein of relative molecular mass 20,000 in the olfactory mucosa (J. Pevs-

*IX Symposium on Olfaction and Taste and 8th Annual Meeting of the Association for Chemoreception Sciences Snowmass Village, Colorado, 20–24 July 1986.

ner, Johns Hopkins University) binds several types of odour molecules with micromolar and millimolar affinities, is localized in olfactory glands and is released into the surface mucus when the glands are stimulated pharmacologically. The protein may facilitate diffusion of odour molecules through the mucus to receptor sites, or it may inactivate and clear odorants from the active sites. Odour molecules also stimulate secretion from olfactory glands by a cholinergic sensitive pathway, which may provide another mechanism by which odour molecules are cleared from olfactory receptor neurones (M.L. Getchell, Wayne State University).

Immunocytochemical techniques can help to elucidate the location of antigenic molecules in olfactory receptor neurones, which will help further studies on the coding of odour quality and the developmental sequences in neuronal maturation. A neurone-specific antigen (RB-8) in the rat has been characterized as a membrane-associated protein with a relative molecular mass of 125,000 (J.E. Schwob and D.I. Gottlieb, Washington University). In contrast to other regions of the nervous system, RB-8 stains the receptor neurone population nonhomogeneously: the axons of receptor neurones in the ventro-lateral region of the olfactory epithelium to the olfactory bulb are RB-8 positive, whereas axons from neurones with apparently identical morphology in the dorso-medial region of the olfactory epithelium are RB-8 negative. There is heterogeneity in cytochrome oxidase staining of olfactory receptor neurones during the development of normal and odour-deprived neonatal rat pups (P.E. Pedersen, Yale University). The further use of these two techniques should provide a clearer understanding of first principles of the functional organization of the peripheral olfactory system and its primary projections to the olfactory bulb, particularly as it relates to the coding of odour quality. Finally, monoclonal antibodies have been developed as specific probes to identify the emergence of membrane antigens during the development of olfactory receptor neurones (A.I. Farbman, Northwestern University; P.E. Pedersen). Two antibodies, Neu-5 and Neu-9, first bind on embryonic day 13 and stain the developing olfactory nerve bundles as well as the soma and dendrites. Another antibody, Neu-4, binds on embryonic day 14, the earliest time at which an odour-evoked response can be recorded from the olfactory mucosa. Further studies should lead to identification of specific membrane antigens that occur in association with the development of odogen- and voltage-gated channels in receptor cell membranes.

In the taste field, there is particular interest in whether Na^+ channels are involved in transduction mechanisms for

salt taste sensation, mainly deriving from studies where amiloride at least partially blocks the response to NaCl and LiCl but not to other salts.

Amiloride is now thought to block the sodium taste response completely, in contrast to previous reports of only partial blockade. Taste responses to sodium acetate and sodium bicarbonate are suppressed by 95–100 per cent after amiloride application on the tongue, whereas responses to NaCl and NaBr are suppressed by 65–90 per cent (B.K. Formaker and D.L. Hill, University of Toledo). Responses to NH_4Cl and ammonium acetate are not affected. Bretium tosylate, an antifibrillary drug that opens amiloride-sensitive Na^+ channels, poten-

tiates the taste of NaCl and LiCl, but has no effect on KCl or CaCl, when applied to the human tongue. The bretium potentiation is inhibited by amiloride.

Other questions arise, such as why the monoclonal antibody 5B4 stains some, but not all, of the taste-bud cells in a bud. Could the antibody be binding a receptor protein that is present in cells only at certain stages of the turnover cycle? □

Gordon M. Shepherd is in the Section of Neuroanatomy, Yale University School of Medicine, New Haven, Connecticut 06510; Thomas V. Getchell is in the Department of Anatomy, Wayne State University, Detroit, Michigan 48021; Charlotte M. Mistretta is in the Department of Oral Biology, University of Michigan School of Dentistry, Ann Arbor, Michigan 48109, USA.

Geophysics

New radiocarbon dating system

Roy Switsur

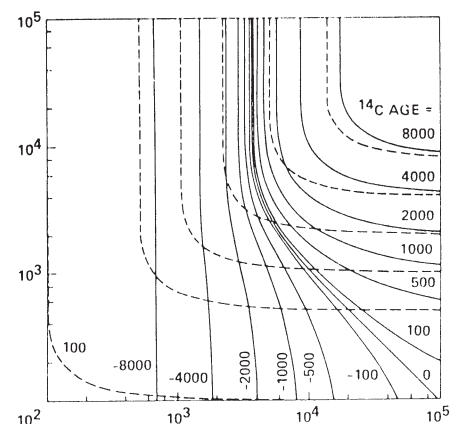
THE closest terrestrial analogue to martian conditions probably occurs in the arid valleys of southern Victoria Land, Antarctica; hypotheses suggesting that life exists on Mars are often based on this comparison. The discovery of cryptoendolithic microorganisms that survive in this inhospitable environment without completely adapting to its extremes add to this speculation. Recent studies of these microorganisms by McKay, Long and Friedmann¹ have led to a new approach to radiocarbon dating which is applicable to systems slowly exchanging carbon with the atmosphere, and which relates the true age to the apparent radiocarbon age and to the carbon-exchange rate.

Dry katabatic winds descend from the Antarctic ice plateau to create a lifeless desert where extensive areas of rock lie devoid of all snow or ice cover. The endolithic microorganisms colonize a narrow zone just beneath the surface of the porous granitic translucent rocks. Because of the hostile environment, photosynthesis can occur during only a few hours each year when conditions of moisture and light are suitable. A detailed analysis of this microbial ecosystem suggests that the organisms have very low rates of metabolism and may be very old².

The method of dating dead organic matter through its residual radiocarbon content, developed 40 years ago by W.F. Libby³, is based on a closed system in which the carbon is unchanged through time apart from that lost by radioactive decay. The initial activity is produced through a chain of natural processes: a constant flux of primary cosmic rays impinging on the outer layers of the Earth's atmosphere generate secondary showers with high neutron content. These neutrons lose energy in elastic collisions as

they descend until, at a height of about 12 km, they react with stable atmospheric ^{14}N nuclei to produce the radioactive ^{14}C isotope $^{14}\text{N}(n,p)^{14}\text{C}$.

Following rapid oxidation the radioactive CO_2 is mixed thoroughly, in a very short time compared with the half-life of the ^{14}C , with the existing CO_2 in the lower atmosphere. Most of the radiocarbon enters the oceans and becomes precipitated



True age in years (vertical scale) against mixing time-constant in years (horizontal scale). Solid line, with bomb effect; broken line, without bomb effect (from ref. 1.)

as a calcium carbonate ooze on the ocean floor. Part of the remainder becomes incorporated into plants through photosynthetic reactions. Of some 80 tons of the ^{14}C isotope that exists on the earth only approximately 3 tons remains in the biospheric reservoir, mainly in the form of photosynthetically produced cellulose. The carbon compounds thus produced by the green leaves of plants form the base of the food net for the animal life of the planet and their natural metabolism results in carbon exchange in equilibrium with the

atmosphere. The ratio of the radioactive to stable carbon in these living organisms is 10^{-12} , producing a specific activity of about 6.8 picocuries per gram of carbon. On the death of the organism, carbon input ceases and the ratio decreases as the ^{14}C decays with a half-life of 5,730 years. Subsequent measurement of the radiocarbon concentration in the organic material permits the elapsed time since death to be calculated by use of the decay equation as long as there is no contamination by carbon from other sources.

It is well known that the assumption of the constancy of the atmospheric activity is erroneous. Measurements⁴ of radiocarbon concentration in samples of growth rings of giant redwood trees of known age show perturbations of 1–2 per cent during the past 1,800 years, and further studies indicate that the natural variation is of the order of 10 per cent. It was realized that a calibration curve is essential for precise radiocarbon ages and much effort has been expended^{5,6} in producing this from dendrochronologically dated Douglas firs and oaks.

Atmospheric activity is also affected by anthropological interventions. The cumulative effect of the combustion of fossil fuels (coal and oil) has diluted the activity to an extent estimated⁷ to be about 3 per cent, so that wood grown, say, in 1950 would appear to be more than 200 years old. This effect is small compared with that of the intense production of radiocarbon by neutrons generated during the detonation of nuclear weapons in test programmes of the late 1950s and mid 1960s. During this short period atmospheric ^{14}C radioactivity increased by 100 per cent with a peak around 1963–64. The subsequent falling off in activity⁸ is crucial to the model of McKay *et al.*

Although this model is applied specifically to the case of radiocarbon, it is capable of wider generalization to other radioisotopes that are diluted by mixing. The basic assumptions are that the carbon leaving the system has the same isotope composition as the total system; that the outward flux of carbon follows first-order reaction kinetics and hence is inversely proportional to the mixing time; and that the ^{14}C is a minute fraction of the total carbon. These assumptions allow a differential equation to be set up for the expression of ^{14}C continuity for a constant volume system and its integration. A further assumption, that the total carbon in the system remains steady over time so that the carbon flux in to the system balances that out, allows the fraction of modern carbon in the system to be estimated. In the absence of perturbations of atmospheric ^{14}C such as those described above, this estimate has a constant value of unity. The nuclear-bomb effect increases this atmospheric ^{14}C content so that the value rises above unity for the period subse-

quent to the commencement of bomb testing. McKay *et al.*¹ fitted a simple exponential curve to Nydal and Lövseth's monitoring data⁸ for the post-1962 decrease in ^{14}C concentration to produce an equation for modern carbon content that does not depend on the true age of the system and may be evaluated for a given calendar date as a function only of the mixing time constant.

The authors also compare the analytical fit with a detailed numerical integration of the bomb radiocarbon measurements for a set of mixing times ranging from 10 to 10,000 years. The larger time-constants give the best agreement. Because the radiocarbon age computation implies a measurement of the ^{14}C content of the system and this is distorted by the temporal variation of modern carbon, the true age differs from the radiocarbon age. It is the value of the mixing time-constant that determines this difference. The figure shows a plot of true age versus mixing time-constant for several values of radiocarbon age, both with and without the bomb effect. The radiocarbon age lies on a locus

of points bounded by the true age.

The problems to which the model might be applied require careful consideration but, apart from dating itself, it could possibly be of use for the study of the exchange rates of carbon in geological or biological systems. It might be used for modelling CO_2 in the atmosphere which requires a knowledge of the exchange rate of carbon between the reservoirs of soil and organic materials. Another intriguing possibility is the study of turn over time of carbon in individual living tissues such as bone collagen or brain. □

1. McKay, C.P., Long, A. & Friedmann, E.I. *J. geophys. Res.* **91**, 3836 (1986).
2. McKay, C.P. & Friedmann, E.I. *Antarct. J. U.S.* **19**, 170 (1984).
3. Libby, W.F. *Phys. Rev.* **69**, 671 (1946).
4. Willis, E.H., Tauber, H. & Münnich, K.O. *Radiocarbon* **2**, 1 (1960).
5. Stuiver, M. *Radiocarbon* **24**, 1 (1982).
6. Pearson, G.W., Pilcher, J.R., Baillie, M.G. & Corbett, D.M. *Radiocarbon* (in the press).
7. Suess, H.E. *Science* **122**, 415 (1955).
8. Nydal, R. & Lövseth, K. *J. geophys. Res.* **88**, 3621 (1983).

Roy Switsur is in the sub-Department of Quaternary Research, University of Cambridge, Cambridge CB2 3RS, UK.

Archaeology

The environmentalist myth

Jared M. Diamond

MANY anthropologists and archaeologists believe that pre-industrial civilizations did far less damage to their environment than do their modern industrial counterparts. Derived from Rousseau's romanticized concept of the 'noble savage', this view holds that hunter-gatherers and neolithic agriculturists feel a reverence for nature, live in harmony with their environment, practise a conservation ethic and avoid the short-sighted destructive exploitation rampant in industrial societies. What reality does this postulated golden age of environmentalism possess? New evidence¹ adds support to the view that the golden age is a myth.

Recent archaeological discoveries have destroyed the assumption that Eden's hunters were conservationists. On every oceanic island for which we have adequate knowledge, the first arrival of humans was quickly followed by the extermination of all or most large animals, the best-known victims being the moas of New Zealand; the giant lemurs and elephant birds of Madagascar; and the flightless geese of Hawaii²⁻⁴. (It is still debated² whether the late Pleistocene extinctions of most large animals of North and South America and Australia were similarly caused by the arrival of humans.) More recently, it has become clear that pre-industrial societies were also prone to habitat destruction on a scale sufficient to destroy them.

One such case involves the collapse of the Polynesian society that erected the famous giant statues of Easter Island. This collapse proves to have been at least partly caused by deforestation, documented by pollen records^{5,6}. When Polynesians reached the island around AD 400, they found it covered with palms, other trees and shrubs. By AD 1500 the human population had grown to about 7,000 and had

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Barren grasslands and the famous monolithic statues of Easter Island. The scene was formerly a palm forest. (Courtesy of Jo Anne Van Tilburg and David Ochsner.)

erected some 245 stone platforms, carved more than 800 giant statues and destroyed the forest so completely that its tree species are now extinct⁷. Deforestation had three harmful consequences: soil erosion, hence lower crop yields; no trees to build canoes, hence lower yields of fish, the main protein source on Easter Island other than chickens; and inability to erect statues (weighing up to 85 tons), which had probably been raised to an upright position with the help of logs used as levers. The human population now exceeded the carrying capacity of the island. Warfare, cannibalism and slavery became chronic, a warrior class took over, spear-points were manufactured in enormous quantities and people reverted to living in caves for defence. By the time that Europeans discovered Easter Island, they found a panorama of destruction: the population may have crashed, many of the statues had been toppled and intentionally decapitated during battles between rival factions, and the island had been converted to barren treeless grassland.

A second example is the collapse of the Anasazi, one of the most advanced pre-Columbian civilizations in the United States. When European explorers reached the deserts of the American south-west, they found gigantic, uninhabited, multi-storied communal houses or pueblos. The dwellings in Chaco Canyon, for example, contained up to 5 stories and 650 rooms, were constructed with more than 200,000 roof beams weighing an average of 275 kg and were the largest and highest buildings in the United States until the advent of skyscrapers in the late nineteenth century. Archaeologists have shown that the dwellings were built about 1,000 years ago and abandoned in the twelfth century. The prevalent view, based on tree-ring evidence for a period of low summer rainfall, is that abandonment was forced by drought. But although drought may have contributed, the Anasazi themselves also inflicted fatal blows on the fragile desert environment.

Betancourt and Van Devender⁸⁻¹¹ used plant remains in radiocarbon-dated pack-rat middens to establish Chaco's vegetational history. When construction of the dwellings began in the early tenth century, the cliffs were covered with pinyon/juniper woodland, which was then cleared for firewood over a radius of 20 km during the next 200 years. The logs to construct the dwellings were mainly ponderosa pine, hauled in via roads from sites mostly more than 40 km away. As pines became depleted, the Anasazi began in AD 1030 to haul spruce, fir and Douglas fir logs from distances exceeding 75 km. In addition, Anasazi desert agriculture depended on gravity-flow irrigation without pumps, so that irrigation would have become impossible if channel erosion eventually lowered the water level below the field

level. With the need to scour over a growing radius for construction timber and fuel, as well as the possible collapse of the irrigation system, the Chaco dwellings had to be abandoned. At Chaco Canyon, as on Easter Island, the environmental damage was irreversible.

Other cases deserve brief mention. It has long been debated whether overgrazing, removal of plant cover and soil erosion contributed to the decline of the ancient civilizations of Greece and the Middle East. A recent study¹ of the Peloponnisos suggests an answer in the affirmative by documenting at least four cycles of settlement, erosion and abandonment between 2000 BC and the Middle Ages. Abandonment was caused by slope erosion and valley siltation, resulting from shortened fallow cycles, clearing of steep slopes, domestic grazing animals and neglect of terracing and check dams. In Polynesia there was massive forest destruction in New Zealand² and Hawaii; degradation of forest to extensive fern savanna useless for agriculture in the Cook, Society, Marquesas and Mangaréva islands; and destruction of the xeric vegetation of Kahoolawe Island, causing a steep population decline and abandonment of all but a few coastal settlements after the sixteenth century³. Many other possible cases, such as the collapse of Mayan civilization and the desertification of the Sahara, cannot be evaluated without detailed evidence.

Although the lack of such evidence until recently surely helped the myth of an environmental golden age to persist, there may be another reason for the myth. It used to be common for industrial societies to treat pre-industrial peoples as savage brutes, meriting extermination or at least dispossession of their lands. In reaction to this appalling history, others have idealized pre-industrial peoples as gentle conservationists. Both views are equally false. Ever since humans acquired the capacity to affect their environment, all societies have faced similar problems of managing fragile natural resources and have risked self-inflicted tragedies. □

1. Van Andel, T. H. *et al.* *Hesperia* **55**, 103 (1986).
2. Martin, P. S. & Klein, R. G. (eds) *Quaternary Extinctions* (University of Arizona Press, Tucson, 1984).
3. Diamond, J. M. *Nature News and Views* **298**, 787 (1982).
4. Steadman, D. W. & Olson, S. L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6191 (1985).
5. Kirch, P. V. *The Evolution of the Polynesian Chiefdoms* (Cambridge University Press, 1984).
6. Flenley, J. R. & King, S. M. *Nature* **307**, 47 (1984).
7. Van Tilburg, J. thesis, Univ. Calif. Los Angeles (1986).
8. Betancourt, J. L. & Van Devender, T. R. *Science* **214**, 656 (1981).
9. Samuels, N. L. & Betancourt, J. L. *Envir. Manag.* **6**, 505 (1982).
10. Betancourt, J. L. *et al.* in *Chaco Canyon Country, a Field Guide* (eds Wells, S. G. *et al.*) 207 (Geomorphological Field Group, 1983).
11. Betancourt, J. L. *et al.* *Am. Antiquity* **51**, 370 (1986).
12. McGlone, M. S. *Archaeol. Oceania* **18**, 11 (1983).

Jared M. Diamond is Professor of Physiology at the University of California School of Medicine, Los Angeles, California 90024, USA.



100 years ago

Observations on heredity in cats with an abnormal number of toes

THIS observation is concerned with a family of four male tabby kittens, all of which possessed the abnormality to a very marked extent. This was the first family produced by a female tabby cat which, when born, was the most abnormal form which had come under my notice, possessing two extra toes on all the paws, i.e. seven on

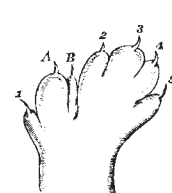


FIG. 1.—Right fore-paw from above, with extra toes.



FIG. 2.—Right fore-paw from below, with extra toes.



FIG. 3.—Right fore-paw from above, normal.

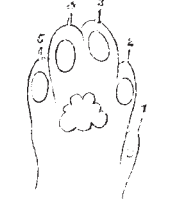


FIG. 4.—Right fore-paw from below, normal.



FIG. 5.—Right hind-paw from above, with extra toes.



FIG. 6.—Right hind-paw from below, with extra toes.

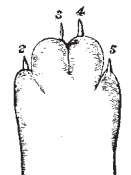


FIG. 7.—Right hind-paw from above, normal.

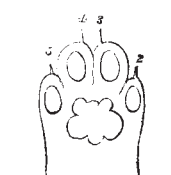


FIG. 8.—Right hind-paw from below, normal.

each fore-paw and six on each hind-paw. The extra toes (in the fore-feet) are those labelled A and B (in Figs. 1 and 2), and they confer the extraordinary breadth upon the foot. The most recently added is B, which is still partially coalesced with A, and has but one pad in common with it (Fig. 2)...There is seen to be an extra pad behind the additional toes, of which there is no trace in the normal foot." In the hind-paws (Figs. 5 and 6) "there is little doubt that the innermost toe I is the hallux lost in the normal foot... The second extra toe is that labelled A... On the under side (Fig. 6) all the toes have separate pads, and there is an additional pad behind the extra toes," which is sometimes fused with that behind the normal toes.

From *Nature* **35**, 38; 11 November 1886.

Where does the message begin?

SIR—The remarkable accomplishment of cloning and sequencing the gene responsible for the X-linked form of the phagocytic disorder chronic granulomatous disease (X-CGD) on the basis of its chromosomal location alone, without reference to any protein, has recently been reported in these pages by Royer-Pokora *et al.*¹ With the amino acid sequence predicted from the gene sequence, it should now be possible to identify and characterize the previously unknown protein responsible for this disorder. However, it seems possible that the true amino terminus of the protein has not been identified.

Royer-Pokora *et al.* conclude that translation of the X-CGD messenger RNA probably initiates at the second in-frame AUG codon, at position 322, because it occurs in a sequence that more closely resembles the consensus sequence for functional initiator codons, defined by Kozak² on the basis of observed sequences, than does that at the first AUG at position 208. The primary difference between the two is that the AUG at position 322 has an A nucleotide three positions upstream, as do 79 per cent of functional initiation codons, whereas at the first AUG there is a U, which is hardly ever observed at this position². On the other hand, 95 per cent of known messenger RNA molecules have translation initiated at the most proximal AUG codon², and the importance of the most proximal position has been demonstrated directly³. The basis of the nucleotide preference observed three nucleotides upstream is not known. Therefore, the choice of neither initiation codon is compelling from just these considerations.

The reason for suggesting that translation of the X-CGD messenger RNA initiates at the first AUG is that the 38 extra amino acid residues indicated by the messenger RNA sequence are not random. In particular, the first 20 residues closely resemble the signal peptides that have been implicated in the co-translational passage of proteins into the endoplasmic reticulum as the first step in targeting the protein for secretion from the cell, insertion into membranes, or sequestration into at least some cytoplasmic compartments⁴⁻⁶. The pertinent properties are that this sequence;

Met-Leu-Ile-Leu-Leu-Pro-Val-Cys-Arg-Asn-Leu-Leu-Ser-Phe-Leu-Arg-Gly-Ser-Ser-Ala-

is positively-charged, has predominantly hydrophobic amino acids from residues 2 to 15, and ends with five polar residues typical of signal peptide cleavage sites. The only unfavourable aspects of the

sequence are the Arg and Asn residues at positions 9 and 10, respectively, within the hydrophobic segment, but such residues are observed at low frequencies in similar positions in known signal peptides⁷. The predictive scheme of Von Heijne⁷ considers the frequencies with which the 12 amino acids proximal to the signal peptide cleavage site, plus the two distal, are observed in known signal sequences. The relative likelihood of the first 20 residues constituting a signal peptide, with cleavage after Ala 20, was calculated to be +5.6, which is well within the range for known signal peptides and outside that found for the terminal amino-acid sequences of cytoplasmic proteins⁷. The first eight residues are not considered in this calculation but are very like those in known signal peptides.

The nature of the X-CGD disorder is consistent with the protein responsible being situated in granules of phagocytes⁸. Proteins designated for such cell sites are often synthesized with signal sequences and are directed to the appropriate cell compartment after entering the endoplasmic reticulum⁹⁻¹¹.

THOMAS E. CREIGHTON

Medical Research Council,
Laboratory of Molecular Biology,
Cambridge, CB2 2QH, UK

1. Royer-Pokora, B. *et al.* *Nature* **322**, 32-38 (1986).
2. Kozak, M. *Nucleic Acids Res.* **12**, 857-872 (1984).
3. Sherman, F., Stewart, S.W. & Schweingruber, A.M. *Cell* **20**, 215-222 (1980).
4. Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1496-1500 (1980).
5. Perlman, D. & Halvorson, H.O. *J. molec. Biol.* **167**, 391-409 (1983).
6. Von Heijne, G. *J. molec. Biol.* **184**, 99-105 (1985).
7. Von Heijne, G. *Nucleic Acids Res.* **14**, 4683-4690 (1986).
8. Tauber, A.I., Borregaard, N., Simons, E. & Wright, J. *Medicine* **62**, 286-309 (1983).
9. Faust, P.L., Kornfeld, S. & Chirgwin, J.M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4910-4914 (1985).
10. Shibahara, S. *et al.* *Nucleic Acids Res.* **14**, 2413-2417 (1986).
11. Lobe, C. *et al.* *Science* **232**, 858-861 (1986).

SIR—We appreciate the comments and suggestions from Creighton¹ and from Ashworth *et al.*² regarding our predictions for the protein product encoded by the gene responsible for X-linked chronic granulomatous disease³.

Creighton suggests that a non-consensus ATG at position 208 of our cDNA sequence might be the correct translation initiation site (rather than the consensus ATG at position 322 we favoured) and that the additional N-terminal extension encodes a signal peptide based on a predictive scheme of von Heijne⁴. As stated in our paper, no certain choice between these ATGs can be made from cDNA sequence alone. Our analyses of the sequence, albeit without the recent approach of von Heijne, led us to favour the ATG at 322. Creighton's proposal needs to be very carefully considered, but in light of our recent finding (R. Brown and S.H.O., unpublished data) that an

intron interrupts the alanine codon he suggests as a signal peptide cleavage site.

Whatever the final resolution, the uncertainty in ATG assignment does not materially affect experimental approaches to identification of the protein *in vitro*. For one, the size difference between the proteins predicted from initiation at either ATG (minus the signal peptide if ATG 208 were used) is small. In addition, we are obliged to examine all cellular fractions to localize the protein to a specific compartment. Finally, initial results with antisera directed against a portion of the predicted amino acid sequences identify an *in vivo* protein located in membrane-rich fractions of granulocytes (M. Dinanuer and S.H.O. unpublished data). Expression of transferred cDNAs will permit definitive assignment of the initiator codon and the possible role of the signal peptide proposed by Creighton in the targeting of the protein within the cell.

Although the suggestion of Ashworth *et al.*² that a region near the C-terminus of the predicted protein resembles a haem-binding site of a cytochrome P450 is intriguing, the significance of the limited resemblance they note is unclear.

STUART H. ORKIN

Division of Hematology, Children's
Hospital,
Howard Hughes Medical Institute,
Harvard Medical School,
Boston, Massachusetts 02115 USA

1. Creighton, T. *Nature* **324**, 21 (1980).
2. Ashworth, A., Shephard, E.A. & Phillips, I.R. *Nature* **322**, 599 (1986).
3. Royer-Pokora, B. *et al.* *Nature* **322**, 32-38 (1986).
4. Von Heijne, G. *Nucleic Acids Res.* **14**, 4683-4690 (1986).

Occam and mankind's genetic bottleneck

SIR—In defending their support for and amplification of the view¹ that non-African human populations are descended from African populations via an emigrant stock that was of very small size, Jones and Rouhani² have exposed the weakness of their argument. It is true that such a bottleneck is an inevitable implication of their simple model. However, that model assumes that the β -globin non-coding sequences of DNA on which their argument is based are unaffected by selection, either directly or by linkage ('hitch-hiking'). This assumption is defended on two grounds. The first is that neutrality of non-coding sequences is implicit in most hypotheses of molecular evolution. But there is actually no evidence for this fashionable assumption. It is adopted merely because of the prejudices of those advancing the hypotheses, because it allows phylogenetic conclusions to be reached in molecular biology laboratories rather than at palaeontologists' benches, and possibly because it makes the models mathematically easier.

Jones and Rouhani tell us that their second reason for assuming neutrality is that, since there is no information on any selective forces affecting the variants they discuss, they "have, with Occam, chosen the more straightforward alternative"². Let us not argue fruitlessly about what William of Occam really meant. In practice, science progresses by attempting to explain phenomena simply, by trying to simplify further, and by adding complexity only when necessary. But what is simpler in this case — assuming no selection and a bottleneck, or assuming selection and no bottleneck? Conceptually the two hypotheses are equally simple: the maths can be simplified either by leaving out selection or by assuming the population size to be effectively infinite. Indeed, the available data are consistent with innumerable combinations of values of the relevant variables — population number, selective coefficients, time and original allele-frequencies. One can, of course, reach conclusions about the values of some of them by making arbitrary assumptions about the values of others but there is no justification for doing so. We just have to admit our ignorance.

Jones and Rouhani are prepared to invoke selection when they need it: they argue against parallel evolution in various populations by recalling Sewall Wright's calculations of the unlikelihood of a chromosome rearrangement becoming established in a population if it is at a selective disadvantage in the heterozygous condition. Do they consider the β -globin sequences to be neutral or do they not?

JEREMY J.D. GREENWOOD

Department of Biological Sciences,
The University of Dundee,
Dundee, DD1 4HN, UK

1. Wainscot, J.S. *et al.* *Nature* **319**, 491-493 (1986).

2. Jones, J.S. & Rouhani, S. *Nature* **322**, 599-600 (1986).

Chernobyl fallout in size-fractionated aerosol

SIR—The Chernobyl accident first produced elevated radioactivity in ambient air in Switzerland on 29 April. Daily high-volume samples showed that the highest activities occurred on 1 May and had already decreased by two orders of magnitude by 8 May. The maximum values observed of ^{137}Cs (2 Bq m^{-3}) are about half of those reported from Studsvik¹.

To understand the interaction of these airborne fission nuclides with the environment, we need to know the carrier responsible for their long-range transport. We therefore measured activity distributions on size fractionated aerosol samples for typical fission products such as ^{131}I , ^{103}Ru , ^{132}Te and ^{137}Cs , using two different commercially available cascade impactors.

From 30 April to 13 May, an Andersen sampler collected aerosols in Spiez (total

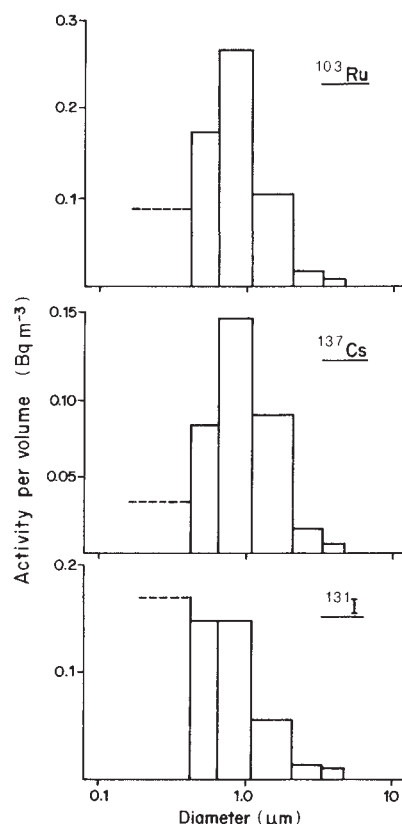


Fig. 1 Activities of ^{103}Ru , ^{137}Cs , and ^{131}I in size fractionated aerosol samples in Bq m^{-3} sampled at Spiez from 09:00 h April 30 to 09:00 h May 13 using an Andersen sampler. Note that the low activities are due to averaging over the 14 day sampling period.

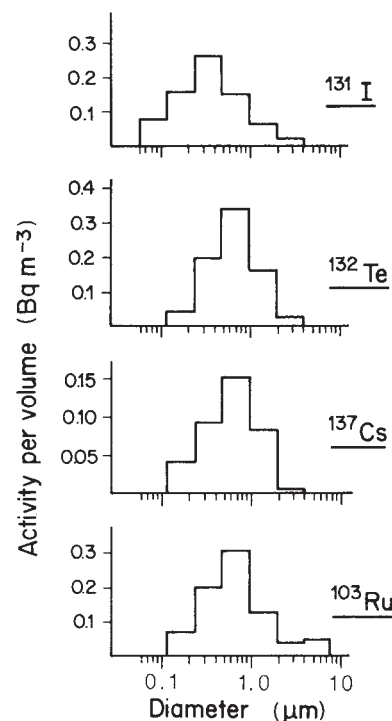


Fig. 2 Activities of ^{131}I , ^{132}Te , ^{137}Cs , and ^{103}Ru in size-fractionated aerosol samples in Bq m^{-3} sampled at Zurich-Hoenggerberg from 09:15 h—May 23:15 h May 6th using a Berner cascade impactor.

570 m^3) and, from 2 May to 8 May, a low pressure Berner impactor² was intermittently in operation in Zurich (total 170 m^3). The impactors fractionate airborne particles into different size classes.

As Fig. 1 shows, ^{131}I has a rather different distribution from the other nuclides. Most ^{131}I is found in the size fraction on the backup filter ($<0.47 \mu\text{m}$), whereas ^{103}Ru and ^{137}Cs show a pronounced maximum at $0.93 \mu\text{m}$ (geometric mean diameter). The Berner impactor (Fig. 2), gives a similar distribution with the maximum for iodine at $0.35 \mu\text{m}$ and at $0.71 \mu\text{m}$ for ^{132}Te , ^{137}Cs and ^{103}Ru . Different distributions are due to the different cutoff diameters of the devices. The distributions for ^{137}Cs , ^{132}Te and ^{103}Ru are very similar to the concentration patterns of prominent ions such as NH_4^+ or SO_4^{2-} and NO_3^- as determined by ion chromatography (data not shown).

Measurements in Sweden¹, Finland³, and Germany⁴ have shown that ^{131}I was mainly ($\sim 80\%$) present in the gas phase. The adsorption of gaseous species onto aerosols is strongly surface correlated^{5,6}; measured surface distributions for environmental aerosols usually peak at size fractions with particle diameters $\leq 0.4 \mu\text{m}$ (refs 7,8). We therefore conclude that the size distribution observed for ^{131}I does not reflect the primary particles released at Chernobyl but rather the surface distribution of the aerosols at the sampling site.

By contrast, ^{137}Cs , ^{131}Te and ^{103}Ru were presumably ejected as particles or attached very early to aerosols and grew by coagulation with other particles during transport. It is therefore reasonable that their size distribution resembles that of the main inorganic ions (sulphate, nitrate and ammonium) which may also be transported over long distances⁹.

In contrast to the findings of Bondietti and Brantley¹⁰, we found no increase of the activity median aerodynamic diameter during the measurement period. We think that this increase may not reflect the original release at Chernobyl, but be due to transport effects. Note that the size distribution of ^{137}Cs , for example, from the Chernobyl fallout is very similar to the size distribution in the fallout of nuclear weapons tests¹¹, in contrast to the size distribution found in the stratosphere, which is shifted towards smaller particles¹².

We have also measured fission products in sequential samples of the first rain at Zurich after 29 April. The activity distributions of the radionuclides were very similar to the concentration distributions of the ions, from which we conclude that the depletion mechanisms are quite similar for both groups. The activity ratio of ^{131}I to the other radioisotopes, such as $^{131}\text{I}/^{103}\text{Ru}$ in rain is about 5, whereas in the impactor samples it is near to one (integrated values from Fig. 1 or 2). This observation also indicates that in the

atmosphere ^{131}I was mainly present as gaseous species and is depleted with high efficiency by rain.

D.T. JOST
H.W. GÄGGLER
U. BALTENSPERGER

Swiss Federal Institute for Reactor
Research

CH-5303 Wuerenlingen, Switzerland

B. ZINDER

Atmospheric Physics ETH
CH-8093 Zurich, Switzerland

P. HALLER

AC-Zentrum

CH-3700 Spiez, Switzerland

1. Devell, L. *et al.* *Nature* **321**, 192-193 (1986).
2. Berner, A. & Luerzer, C. *J. Phys. Chem.* **84**, 2079-2083 (1980).
3. Finnish Centre for Radiation and Nuclear Safety *et al.* *Second interim report: Radiation situation in Finland from 5 to 16 May 1986* (STUK-B-VALO 45, Helsinki 1986).
4. *Umweltradioaktivität und Strahlenexposition in Südbayern durch den Tschernobyl-Unfall* (GSF-Bericht 16/86) (Gesellschaft fuer Strahlen- und Umweltforschung, Munich, 1986).
5. Raunemaa, T. & Hautajärvi, A. *J. Aerosol Sci.* **15**, 123-132 (1984).
6. Kruger, J. & Nöthling, J.F. *J. Aerosol Sci.* **10**, 571-579 (1979).
7. Seinfeld, J.H. *Atmospheric Chemistry and Physics of Air Pollution*, 289 (Wiley, New York, 1986).
8. Whitby, K.T. *et al.* *J. Colloid Interface Sci.* **39**, 177-204 (1972).
9. Weiss, R.E. *et al.* *Science* **195**, 979-981 (1977).
10. Bondietti, E.A. & Brantley, J.N. *Nature* **322**, 313-314 (1986).
11. Lockhart, L.B., Jr. *et al.* *J. geophys. Res.* **70**, 6033-6041 (1965).
12. Persson G. & Sisefsky, J. *High Phys.* **21**, 421-428 (1971).

Alimentary restrictions and ^{131}I in human thyroids

SIR—Hill *et al.*¹ have reported measurements of ^{131}I during the period following the Chernobyl reactor accident. We have directly determined ^{131}I activity in specimens of thyroid gland obtained at autopsy from 51 adult individuals dying of accidental causes in the area of Genoa. Sampling started on 3 May 1986, 2 days after arrival of the radioactive plume over the area and one day before an abundant rainfall.

Each thyroid specimen was placed in the central well of a shielded NaI(Tl) detector connected with a 1,024-channel spectrometer; radioactivity was measured against a ^{131}I standard having the same geometry.

Differences between individuals (Fig. 1) presumably reflect variable exposure. High levels of ^{131}I were detected from the beginning of radioactive fall-out. The decay of radioactivity in the following weeks was less pronounced than expected on the basis of the effective half-life of ^{131}I in thyroid (7.55 days) (Fig. 2), inferred from its biological (120 days) and physical (8.05 days) half-life.

Although these data are consistent with some persistence of the exposure to ^{131}I , the progressive decrease of radioactivity in the thyroid contrasts with the initial increase of ^{131}I in milk samples from the same area (Fig. 2), presumably reflecting the efficacy of alimentary restrictions in-

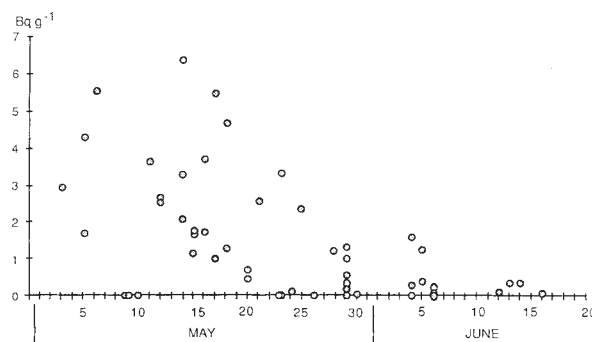


Fig. 1 ^{131}I activity as measured in 51 thyroid specimens in the period following the Chernobyl accident.

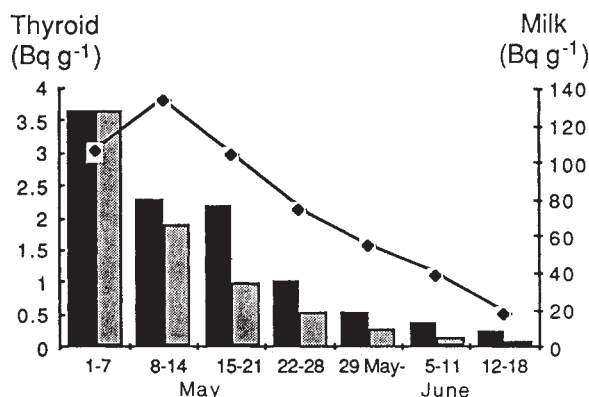


Fig. 2 Mean weekly ^{131}I activity in milk (continuous line) and in thyroid (black columns) during the 7 weeks following arrival of the radioactive plume. Stipled columns indicate the theoretically expected decay of radioactivity, starting from the levels detected in thyroid during the first week of monitoring.

produced in Italy on 3 May. It is also significant that, on withdrawal of restrictions (16 May), the trend towards the loss of radioactivity in the thyroid was temporarily discontinued (Fig. 2).

The mean absorbed β dose was inferred from the ^{131}I activity measured in thyroid specimens, using an *ad hoc* conversion factor². During the first 7 weeks following arrival of the radioactive plume, the mean ($\pm$ SD) weekly absorbed doses were 3.66 ± 1.68 , 2.31 ± 2.12 , 2.20 ± 1.61 , 1.02 ± 1.37 , 0.55 ± 0.58 , 0.38 ± 0.51 , and 0.24 ± 0.17 $\mu\text{Sv g}^{-1}$ respectively. Assuming an average weight of 20 g per thyroid, it can be estimated that the cumulative load absorbed during that period accounted for a mean value of 207 μSv per individual. These estimates are higher than those

made for adult individuals in southern United Kingdom, where the environmental contamination, as deduced from milk ^{131}I activity, was substantially lower than in our region¹.

P. ORLANDO
G. GALLELLI
F. PERDELLI
S. DE FLORA

Institute of Hygiene

R. MALCONTENTI

Institute of Forensic Medicine,
School of Medicine,
University of Genoa,
16132 Genoa, Italy

1. Hill, C.R., Adam, I., Anderson, W., Ott, R.J. & Sowby, F.D. *Nature* **321**, 655-656 (1986).
2. Snyder, W.S., Ford, M.R., Warner, G.G. & Watson, S.B. *MIRD pamphlet No. 11* (Society of Nuclear Medicine, New York, 1969).

Radioactive fallout swipe samples from Chernobyl

SIR—Significant quantities of radioactive fallout from Chernobyl were detected in aircraft swipe samples throughout the month of May, initially on flights from Moscow and, later, from London, New York and Tokyo. Swipe samples have provided valuable information on atmospheric nuclear weapon tests^{1,2} in the past. The fallout from Chernobyl was measured in the swipe samples collected at Bombay from 2 May onwards. Though the radioactive plume rise at the accident site has been estimated³ to be around a kilometre, atmospheric dispersion with time could carry the activity to higher altitudes which enabled its detection.

In all, 20 isotopes; ^{95}Zr , ^{95}Nb , ^{99}Mo ,

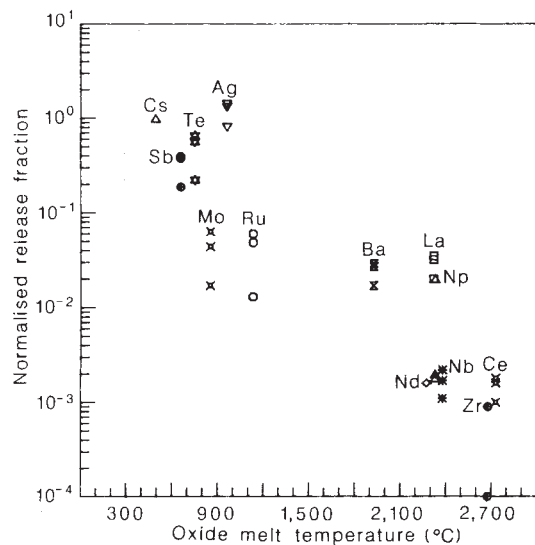
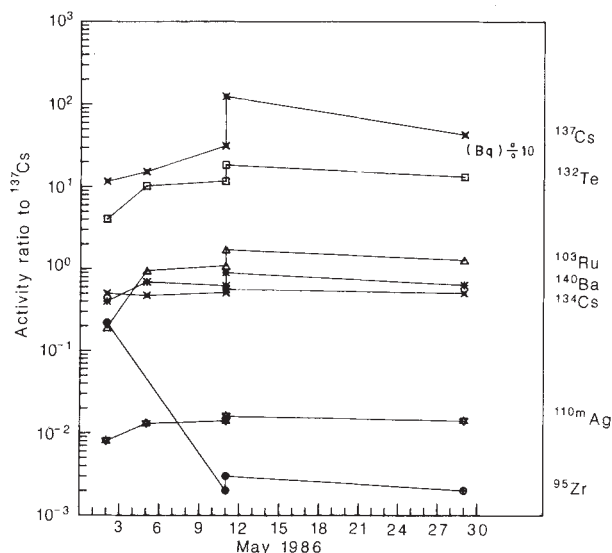
^{103}Ru , ^{106}Ru , ^{110m}Ag , ^{111}Ag , ^{125}Sb , ^{127}Sb , ^{132}Te , ^{131}I , ^{134}Cs , ^{136}Cs , ^{137}Cs , ^{140}Ba , ^{140}La , ^{141}Ce , ^{144}Ce , ^{147}Nd and ^{239}Np were measured by high resolution gamma spectroscopy. Although the nature of the sampling does not permit quantification with respect to volume of air, useful information could be obtained by studying suitable radioisotope ratios.

Our results are generally similar to those reported by others³⁻⁶. Different isotopes of caesium, iodine, ^{132}Te , ^{140}Ba , ^{140}La , ^{103}Ru and ^{99}Mo were the major contributors to the activities measured in the samples. Comparisons with similar swipe samples collected after the last atmospheric test in 1980, which was a low altitude H-bomb explosion, show that in the present case the ^{137}Cs levels are about an order of magnitude higher and ^{103}Ru

Table 1 Activity Ratios (on 26.4.86) of fission products and ^{239}Np to ^{137}Cs in 5 swipe samples

Ratio of element to ^{137}Cs	Origin of flight and the date of collection at Bombay				
	Moscow 2.5.86	New York 5.5.86	New York 11.5.86	Tokyo 11.5.86	Tokyo 29.5.86
^{95}Zr	0.022	—	0.002	0.003	0.002
^{95}Nb	0.05	0.04	0.025	0.026	0.019
^{99}Mo	0.41	1.04	1.48	1.81	—
^{103}Ru	0.19	0.95	1.09	1.72	1.27
^{106}Ru	0.065	0.18	0.24	0.35	0.29
$^{110\text{m}}\text{Ag}$	0.008	0.013	0.014	0.016	0.014
^{111}Ag	0.8	—	—	—	—
^{125}Sb	0.017	0.04	0.03	0.05	0.03
^{127}Sb	0.24	0.43	0.51	0.56	—
^{132}Te	4.05	10.2	11.5	18.4	13.2
^{131}I	1.84	1.06	0.65	2.13	1.13
^{134}Cs	0.50	0.47	0.51	0.56	0.50
^{136}Cs	0.19	0.17	0.20	0.22	0.20
^{140}Ba	0.40	0.69	0.62	0.90	0.63
^{140}La	0.50	0.79	0.87	1.19	0.82
^{141}Ce	0.04	0.02	0.04	0.05	0.05
^{144}Ce	0.07	—	0.02	0.02	0.03
^{147}Nd	0.014	—	—	—	—
^{239}Np	0.42	0.48	4.9	7.1	—
^{137}Cs	116	152	315	1241	427

activity
(becquerels)

**Fig. 1** Activity ratios of selected fission products with respect to ^{137}Cs . The activity of ^{137}Cs in the samples are also shown.**Fig. 2** Release fraction with respect to Cs plotted against fission product oxide's melting point.

an order of magnitude lower. The levels of ^{132}Te and ^{99}Mo , are 10–100 times lower. The other major difference is that the ^{95}Zr concentration is 1000 times less, than from the samples after the last nuclear test.

The isotope activity ratios to ^{137}Cs (as on 26 April) in 5 of the samples collected during May are given in Table 1 and a few selected ratios are plotted in Fig. 1. The ratios of elements whose oxides have high melting temperatures (for example Zr, Nb and Ce) with respect to ^{137}Cs were higher in the first sample collected on 2 May, probably representing the initial release from the reactor. Subsequent samples show higher relative concentrations of Ru, Te, Mo and Sb and, to a lesser extent, of Ba, La and Ag. The $^{131}\text{I}/^{137}\text{Cs}$ ratio variations in these samples did not show any set pattern, perhaps because much of the ^{131}I may be in vapour form⁴.

Two estimates of the reactor fuel burn up have been reported^{4,6}, based on $^{134}\text{Cs}/^{137}\text{Cs}$ and $^{103}\text{Ru}/^{106}\text{Ru}$ ratio values of 0.5 and 4.5 respectively; we observed similar values for the ratios. We also obtained other ratios from the isotope pairs of the same elements such as $^{127}\text{Sb}/^{125}\text{Sb}$, $^{136}\text{Cs}/^{134}\text{Cs}$ and $^{141}\text{Ce}/^{144}\text{Ce}$, the last of which varied somewhat from sample to sample. We obtained reactor core inventories from the ORIGIN 2 code⁷ for 1.8% enriched uranium Candu type reactor and compared them with the observed activity ratios. They correspond to a fuel exposure of around 650 d (700 d if 2.0% enriched U was used) though our $^{134}\text{Cs}/^{137}\text{Cs}$ also gave fuel exposure of 400d.

The release fractions, normalized to ^{137}Cs , vary with the melting point of the oxides of the fission products for the first three samples. The fractions for Sb and Te agree well with those documented for the PWR reactor core melt accident⁸ while for other elements marked differences are observed. Deviations start from Mo onwards and increase with increasing oxide melting point, suggesting that there may have been only a partial core melt (Fig. 2).

We thank V.K. Sundaram, Health Physics Division, for the ORIGIN 2 results and Dr V.K. Shukla, G.N. Shaikh and T.N. Mahadevan for help in sample collection.

S. SADASIVAN
U.C. MISHRA

Air Monitoring Section,
Bhabha Atomic Research Centre,
Bombay 400 085, India

- Mishra *et al.* *J. sci. indus. Res.* **33**, 216–221 (1974).
- Sadasivan, S. & Mishra, U.C. *J. geophys. Res.* **87**, 7343–7345 (1982).
- Interim report on radiation situation in Finland from April 26 to May 4, (STUK-B-VALO 44) (Finnish Centre for Radiation and Nuclear Safety, Helsinki, 1986).
- Devell *et al.* *Nature* **321**, 192–193 (1986).
- Fry, F.A., Clarke, R.H. & O'Riordan, M.C. *Nature* **321**, 193–195 (1986).
- Hohenemser *et al.* *Nature* **321**, 817 (1986).
- Croff, A.G. *Nucl. Tech.* **62**, 335–352 (1983).
- U.S. Nuclear Regulatory Commission Reactor Safety Study (WASH 1400, NUREG 75/014) (Washington, 1975).

Pragmatism in public health

Alan Glynn

Inventing the NIH: Federal Biomedical Research Policy, 1887–1937.

By Victoria A. Harden.

The Johns Hopkins University Press: 1986. Pp. 274. \$32.50, £25.50.

DAVID Smyth, the physiologist, once said "If primitive man had something on his mind it was probably because he had nothing in his stomach". The primitive hunger which led eventually to the founding of the National Institute of Health was simply for public health. The NIH's forerunner was the Laboratory of Hygiene set up in 1878 on Staten Island, New York, to detect infections among immigrants by using the new science of bacteriology. The Laboratory was part of the Marine Hospital Service, itself founded in 1798 to provide medical care for seamen.

Victoria Harden has drawn on numerous primary and secondary sources for her analysis of the forces which led to the establishment of the NIH in 1930. The author is not a scientist and her book is not a history of bench science or laboratory medicine. It is a history of the pressures, scientific and political, private and public, which shaped United States federal policy on biomedical research. In the end it was this policy which, after 50 years of struggle, allowed the change from a one-man, one-room laboratory to the first National Institute of Health. Development of federal policies over 50 years produced eleven Institutes spread over 317 acres at Bethesda in Maryland, employing 14,000 people and using or distributing 30 per cent of all American health research funds.

There was a false start early on. The value of "sanitary science" had become apparent during the Civil War and after

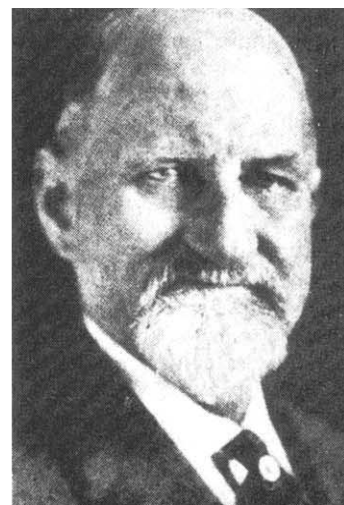
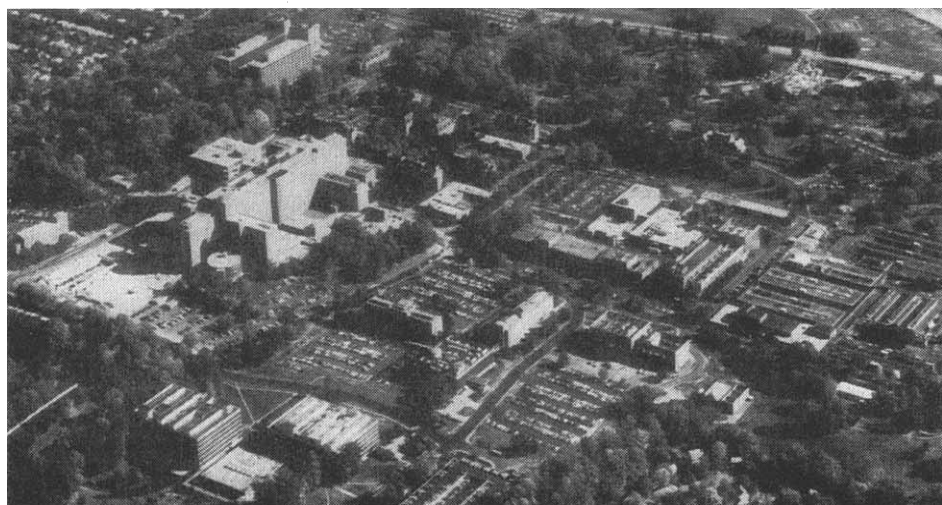
the War attempts were made to enlarge and reform the Marine Hospital Service by, among others, John Shaw Billings, now best known as the initiator of the *Index Medicus*. Billings failed at first but an epidemic of yellow fever in the Mississippi valley led to the formation of a new National Health Board with him in charge. The Marine Hospital Service fought back and, starved of funds, the Board soon withered away. This set a recurring pattern. Reformers, and there were many, came up against not only the natural inertia of a conservative Congress or President but the opposition, open or covert, of the grandees of the Marine Hospital Service who were determined that whatever change had to come they were going to remain in charge.

Harden has, rightly, followed through particular lines of development rather than given a strictly chronological account. There is a list of the principal advances made by workers in the Laboratory of Hygiene, but for those who lack a detailed knowledge of American history a table of the main political events would have been more useful. The Laboratory flourished during the "progressive era" before the First World War. The War, and the work it demanded, stimulated scientists to think of major collaborative research projects embracing many disciplines, though there was still an instinctive distrust of the institutionalization of science.

The book outlines the uneasy and shifting alliances between scientists, health workers and reforming politicians, and helps us to understand an apparent paradox: how a country with a strong philosophy of self-help and freedom from government interference managed to create the largest and most successful laboratory ever devoted to biomedical research almost entirely on government money. Except for one or two brief references, for Harden Europe does not seem to exist. This is a pity, because it is fascinating to compare the events and philosophies she discusses with what was happening at roughly the same time on the other side of the Atlantic.

For example, in its early days the Laboratory of Hygiene was only authorized to deal with practical matters of infectious disease. Its early researches into pellagra or even water pollution were, strictly speaking, illegal uses of government money; indeed a separate appropriation for research was not made until 1907. Sir John Simon in England had faced similar problems 40 years before, when he imposed his own ideas on a vague remit from the Privy Council and eventually persuaded them to authorize, as he put it, "investigations, not necessarily connected with our practical business at the moment . . .". That was in 1865, and because governments everywhere "respond to pestilences" Simon was no doubt helped by the spread of cholera and cerebro-spinal fever in Europe and the arrival of yellow fever in Swansea. As a distinguished pathologist and surgeon, Simon received much help and backing from his professional colleagues whereas the early proponents of the NIH were sometimes in conflict with the American Medical Association.

Perhaps the closest parallel in Britain is the setting up of the Medical Research



The National Institutes of Health in the 1980s, showing only part of the Bethesda complex. On the right is Senator Joseph E. Ransdell, sponsor of the bill that transformed the Laboratory of Hygiene into the NIH. (Pictures respectively courtesy of the NIH and from the book under review.)

Council, the origins of which go back to the Local Government Board of 1871 and more immediately to the Medical Research Committee of 1913. Perhaps because of the tradition of Chadwick, Simon and Farr there seems to have been less worry in Britain than in the United States about the state getting mixed up in health and research, and the MRC was protected from direct ministerial interference at least until the Rothschild era. Although the NIH also achieved considerable scientific independence, Congress continues to try to influence its projects — as in the bill to set up an Institute of Nursing, vetoed last year by the President.

Despite the prosperity of the 1920s, progress in public health was difficult under the "new era" politics of Coolidge. Coolidge never mentioned the subject in a speech, and though treasury secretary Mellon was more favourably inclined he could rely on the director of the budget, Herbert Lord, to keep the purse strings tightly drawn. Lord combined a professional dislike of spending money, especially on research, with the Christian Scientist's belief in its futility.

By the late 1920s, the story takes on some of the inevitability of a Greek tragedy. Harden describes in detail the negotiations through which, after many years of struggle, Congress was persuaded to pass the Parker bill giving more freedom to the Public Health Service. Eight days later Coolidge vetoed it. Again there is a detailed but lively account of the haggling and manoeuvring in Congress. This time the Ransdell bill, which specifically set up the National Institute of Health, failed partly because of opposition from Parker who was smarting at his own defeat. At last Coolidge went and Lord retired. After more hectic activity, both bills passed both houses and were signed by President Hoover. It was 1930 and the Great Depression had already started.

Throughout this story there are many colourful and interesting characters. Yet I have the impression that the colour and the passion have been carefully filtered out in the interests of analytical clarity. This is unfortunate for the amateur historian or the general reader, but for those who want to influence public health policy the message comes through clearly. They must educate the public, the legislators and the administrators. And education here means both short- and long-term campaigns. It means using every effort to win over pressure groups, influencing the influential and converting the doubtful. Well-timed epidemics are always a help. In short this is an excellent, though rather dry, textbook for wets. □

Alan Glynn is Director of the Central Public Health Laboratory, Colindale, London NW9 5HT, UK, and a Visiting Professor at the London School of Hygiene and Tropical Medicine, University of London.

Signal studies in psychology

D. Aled Jeffreys

Human Brainwaves: The Psychological Significance of the Electroencephalogram.

By Jacob Empson. Macmillan, London/Stockton, New York:1986. Pp.126. Hbk £25, \$50; pbk £7.95.

In 1934, Adrian and Mathews confirmed and extended Berger's earlier but largely ignored discovery that small, continually varying, electrical signals could be recorded from the human scalp. These electroencephalograms (EEGs) have been extensively studied over the past 50 years but, even so, relatively little is known about their origins and significance. In contrast to their well-proven clinical applications, therefore, the "spontaneous" EEG signals have had only limited uses in research into normal brain function. They have nevertheless proved to be useful indicators of the level of consciousness and, in particular, of the different stages of sleep. There have also been encouraging results from experiments in which stimulus-induced changes in the EEG — the so-called event-related or evoked potentials (ERPs or EPs) — have been used to investigate cognitive and, to a lesser extent, sensory aspects of brain function.

Newcomers to such electrophysiological studies of the human brain are unfortunately confronted with a vast and specialized literature, often with a high technical content. Jacob Empson has tried to alleviate the problem by providing a simple, non-technical introduction to the use of the EEG and ERPs in cognitive psychology. The first two chapters include brief accounts of the history of the EEG and of modern encephalographic practice;

descriptions of the typical characteristics of the EEG and their use in the assessment of states of consciousness; and discussion of various theories on the origins of the alpha rhythm. There are also individual chapters on event-related (and slow) potential studies of psychological processes, and on the EEG and localization of function in the brain. Two final chapters are devoted to the author's particular interests of normal and abnormal aspects of sleep.

As expected, the book does not provide detailed reviews of the literature but instead contains selective summaries of what the author judges to be the most important experimental findings. One can, in several cases, argue with some of his particular choices or omissions; and also perhaps complain that a short summary similar to that given in the section on event-related potentials was not provided in every chapter. But such shortcomings must be judged against the difficulty of the author's task. The quality of production is quite good except that the print on several pages of my review copy was slightly smudged.

The book is written in a simple and clear style, although occasional sentences have uncertain meaning and, in two cases (pp. 17 and 31), do not make sense at all, presumably because of missing lines or words. On the whole, however, it is easy to read and meets its main objective of providing a coherent and non-technical survey of the topic. *Human Brainwaves* will serve as a useful primer for psychology students and others who need a general introduction to non-clinical, psychological aspects of the EEG. For those who wish to progress further, each chapter has a selected bibliography of more comprehensive review papers. □

D. Aled Jeffreys is Senior Research Fellow in the Department of Communication and Neuroscience, Keele University, Staffordshire ST5 5BG, UK.

Material success

Charles Newey and colleagues

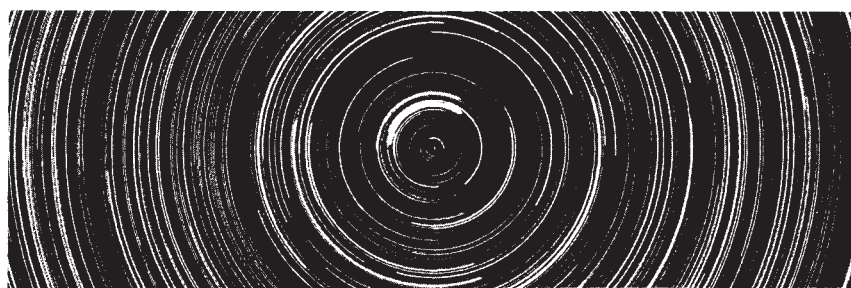
Encyclopedia of Materials Science and Engineering. Editor-in-chief Michael B. Bever. Pergamon:1986. Eight volumes, pp. 6,015. £1,350, \$1,950.

THE *Encyclopedia of Materials Science and Engineering* (EMSE) contains some five million words in seven volumes, plus a volume of indexes. How could a single reviewer possibly do justice to such a massive effort?

Depending on their stance inside or outside the "materials community", different users will look for different things from this encyclopaedia and one person could

not hope to have a representative view. Further, review of EMSE requires assessment of a project planned by about 50 people and implemented by nearly 1,300 contributors writing over 1,500 entries. So, in keeping with the traditions of The Open University, we adopted a team approach. What follows is the essence of our corporate appraisal.

A good test of the usefulness of any encyclopaedia is to see how it performs against a list of topics you would expect or like to see covered. Our list included topics as diverse as Kyropoulos crystal growth, zeolites, charged dislocations, bubble memories, pyrometry, mercury reserves, globular proteins and value analysis. EMSE passed with flying colours, only failing on design for recycling (but recycling is well covered) and charged disloca-



*The Earth moved — star tracks on an 8-hour exposure taken by a fixed camera pointed towards the North Pole; the boldest trail near the centre was made by Polaris. The picture is taken from the new paperback edition of *Orbiting the Sun: Planets and Satellites of the Solar System* by Fred Whipple (reviewed in *Nature* 294, 21; 1981). Publisher is Harvard University Press, price is \$9.25, £6.75.*

tions (hardly surprising). In fact, the coverage is much more diverse than this with treatments of topics not conventionally associated with the books' title, for instance design, safety and health, materials and machines in antiquity, materials economics and management, art materials, newsprint and many materials of biological origin.

The Systematic Outline of EMSE (see Vols 1 and 8) groups the article titles under the 45 main subject headings used in the commissioning; they cover the production of materials, their nature and applications, property and structure assessment techniques, surface phenomena and processes, and more general topics such as mineral resources and management, and economics. These are further subdivided "as each subject dictates". The Introduction states that the space devoted to each of these sub-sections is in relation to its scientific, technical and industrial importance but we are given little insight into how this was assessed. The basis of selection must, in the end, be subjective and the updating policy will show whether the editors have met their goals.

The 1,570 articles are arranged alphabetically, which is a rather mixed blessing. This format should force users to look outside their specialities and, one hopes, help the editors to achieve their aim of cross-fertilization. The juxtaposition of "Titanium Alloys" with "Timbers from Various Geographical Locations" and "The Effect of Tax Upon Materials Policy" demonstrates this amply. On the other hand, the scheme puts a lot of stress on the cross-referencing system and it doesn't hold up too well, especially in terms of alternative titles (for example, "glassy" and "amorphous"). In addition, because the articles are not grouped under, say, specific materials, properties, processes or artefacts, the subject index — which is in Vol. 8 and very thorough — is an absolute must if you are looking for something specific.

Basically, EMSE consists of two types of entry: specific articles on particular topics in the science and engineering of materials, and articles of a more general kind. By and large the specific articles are

authoritative, if succinct, treatments of their subjects but there is considerable variation in the level and approach. Perhaps this is inevitable, given the need to involve so many different authors each with their own ideas about the task, but there does seem to have been a certain unwillingness either to give clear briefing or to wield the editorial pen in the quest for consistency.

It is with the general articles that EMSE is clearly trying to be more than a conventional encyclopaedia. These are either reviews (for example "Electrical Materials — An Overview" and "Energy Considerations in Metals Production") or articles of an integrative nature ("Ceramics: Relation of Microstructure to Processing History", "Product Liability and Materials" and "Design and Materials Selection and Processing", for instance). The integrative articles are particularly informative and stimulating; they are good illustrations of the way to link aspects of materials that are often treated in isolation and, therefore, inadequately. Their inclusion is a clear manifestation of the editors' laudable desire to lower the barriers between disciplines. But beware, you will not find these gems in the subject index — for these you will need to use the Systematic Outline (Vol. 8).

In addition to unevenness in the level of the articles, the work can also be criticized for its slight bias towards the United States, and for a hint of partisanship. However, at this early stage in what should be an evolutionary process for EMSE, we will call this carping and trust that the balance is adjusted in the next edition.

The editors have tackled an enormous challenge and have met with considerable success. These volumes will be an important reference source for all scientists and engineers involved with the properties, processing and uses of materials, and especially for teachers of courses in the materials field. □

Charles Newey, Mark Endean, Peter Lewis, Nick Reid, Graham Weaver, George Weidmann and John Wood are members of the Materials Discipline, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK.

Living for physics

Michael Holwill

Introductory Biophysics. By M. Cerdonio and R.W. Noble. *World Scientific:1986.* Pp.216. £20.25, \$25.30.

FOR those who browse among the shelves of libraries or bookshops, the title of this book and the attractive seascape on the dust-jacket will provide little indication of the contents. It is not an introductory text for students embarking on a course in biophysics; rather, it is aimed at scientists with a reasonable background in physics, particularly in thermodynamics, and it is intended to demonstrate that the basis of living systems can be understood and explained by the application of physical principles. This is not a new concept, but recent developments in both molecular cell biology and irreversible thermodynamics make it appropriate to present a modern account of the ideas concerned.

After a brief description of the characteristics by which living systems can be identified, and a chapter dealing with the structure and organization of those systems, from viroids through bacteria to multicellular organisms, the authors apply physical ideas to a number of biological areas before discussing, in the final chapter, the origins of life. The chapters are essentially independent, each containing some interesting features but having little direct relationship to the others. There is, in fact, much of importance and value in the book, but the rather dated diagrams (which vary considerably in style and quality) and the use of old-fashioned units (for example, "entropy units" instead of $\text{J K}^{-1} \text{mol}^{-1}$ and "calories" for joules) detract from the modern scientific approach.

It is convenient to have, collected together in a single volume, accounts of the essentials of various physico-chemical processes and techniques relevant to a comprehensive examination of the nature and development of living systems. By reducing the space occupied by some of the diagrams, however, a little more introductory material could have been added without the need for additional pages, and this would have increased the appeal of the work to the great number of scientists who have a general interest in this controversial topic. In the final chapter the authors provide some pertinent insights into how thermodynamics, chemical kinetics and fundamental particle physics may be of importance in specifying evolutionary processes, but they have not capitalized on the material earlier in the book by giving a clear summary of the current status of this analytical study. □

Michael Holwill is Reader in Biophysics in the Department of Physics, King's College London (KQC), Strand, London WC2R 2LS, UK.

That certain time in history

W.H. Brock

The Age of Science. By David Knight. Basil Blackwell: 1986. Pp.251. £17.50, \$24.95.

THE reappraisal of the Victorians began in the 1940s. Much derided in the 1920s, today even politicians are reminding a specially-selected set of their "values" for the 1980s. Notwithstanding the material and cultural gaps between 1986 and 1886 or 1836, much of our present world can be clearly seen as derived from the Victorians — notably because of the vigorous export, through travel and emigration, of Victorian institutions and attitudes overseas, including parliamentary government, schools and universities, and the civil service, and religion and attitudes towards women and the family. Arguably, even the harmonies of contemporary popular music are those of Victorian church music. And, outside the relativistic and quantum boundaries of atomic physics and astrophysics, our vaunted twentieth-century science is essentially Victorian in origin.

In *The Age of Science*, David Knight, the most prolific of contemporary English historians of science (this is his eighth book), examines the scientific world-view in the nineteenth century. Although we may well believe that our present era is unprecedented in its domination by science, the reader will be persuaded by

Knight that science equally dominated nineteenth-century thought and debates on religion, education, literature and the economic viability of the nation, though in a very different manner to our own. Here was a period before disciplines became so specialized that their practitioners became alienated from one another, when scientist, clergyman, writer and artist still spoke a common cultural language, and when science was ambitious, cocksure of progress and in its hegemonic role as the Church Scientific. As Knight points out, whereas historians may (wrongly) ignore the impact of science on earlier centuries, science cannot be left out of the narrative of Victorian history or of the historians' repertoire of explanations of what happened in history.

Knight writes deftly and entertainingly. Whether examining Anglo-French scientific rivalry, the challenge of German science, Victorian criteria for science and pseudoscience, the use and significance of scientific illustrations, symbols and jargon, the symbiosis of Victorian science and religion, or the slow erosion of scientific certitude, he wears his learning lightly. *The Age of Science*, a summary of the specialized reading Knight has done during the past 20 years, is a *tour de force*. It not only succeeds in making the history of science popular, but restores its 1950s role as a bridge between the cultures and disciplines which are the consequence of the Victorians' age of science. □

W.H. Brock is Director of the Victorian Studies Centre and Reader in the History of Science at the University of Leicester, Leicester LE1 7RH, UK.

Cold comforts

Beverly Halstead

How to Deep-Freeze a Mammoth. By Björn Kurtén. Columbia University Press: 1986. Pp.121. \$16.96.

A SCIENTIST's general musings in a series of light essays on his subject are the dread of reviewers; few ever rise to the riveting heights of a Stephen Jay Gould. Björn Kurtén's slim volume, however, is well above the average. Kurtén specializes in the study of mammals of the Ice Age, and he has written fictional accounts of the life and times of the period.

Most of the essays deal with aspects of the Pleistocene, but also cover such topics as the third early bird *Archaeopteryx*, the early history of continental drift and the great Messinian salinity crisis when the Mediterranean dried out. This last event came to light when salt deposits 1.5–2 km in depth were discovered during the deep-sea drilling programme; after the drought came the flood, as the Atlantic

poured in through the Straits of Gibraltar — the largest waterfall in Earth history.

It is, however, in writing about the Pleistocene that Kurtén has the surest touch — he does in fact explain how to deep-freeze a mammoth. It has happened but rarely. Normally the outer layers freeze but the deeper parts putrefy and all that remain are bones and bacteria. However, on a southern-facing slope in winter, it is possible for an entire animal to be freeze-dried, and Kurtén describes the discovery of a 30,000-year-old bison from Alaska which made a palatable stew.

There is, too, an essay on Piltdown Man; here, being a Finn Kurtén misses some of the subtlety of English humour. He refers to the "elephant club" from Piltdown that had been carved with a steel knife, but this unique artefact was seen by all, even the duped Sir Arthur Smith-Woodward in his book *The First Englishman*, as akin to a cricket bat. What else would the first Englishman require? The first European in the form of the now-notorious *Petralona* skull from Greece is featured, but in this case the discussion verges on the technical as Kurtén has been

personally involved. In general, though, he has a light touch. Most of the essays are very readable, and the first and last of them are outstanding.

In the first, entitled "The Two Cultures", Kurtén contrasts the fundamental difference between physics and chemistry on the one hand, and biology on the other, expressing with great clarity the attitudes of many practitioners of the "woolly" sciences:

But the world of physics is no more than a corner of the world of science. A man is greater than a star. A bacterium is greater than a star. A star cannot become very much else than a star — but bacteria have evolved into men Biology is the study of individuals, all of them unique, in inexpressible multiplicity and richness. And so the world of biology is something very different from the world of physics and calls for a different approach and different concepts.

The last essay is on cave art. It may be wondered if there could possibly be anything more to say that has not already been said many times before, but one surprising fact is that it appears that the Ice Age was a Golden Age: there was an abundance of food and the lifespan of human beings was not paralleled until virtually the present day. Furthermore, when cave art is examined from the standpoint of a zoologist an unusual perspective emerges: a male reindeer on heat or seals swimming upside down, there is a remarkable precision of observation. Kurtén suggests that many of the paintings may have been for children, in fact for their education. Interestingly, children's footprints are common on the floors of these caves. It would be important for them to recognize the animal's state and hence predict its behaviour.

When it comes to the portrayal of the human form, Kurtén points out that here too there is great accuracy of observation. He notes that many examples of cave art portray sexually inviting attitudes and are not dissimilar to photographs in pornographic magazines of the present day. The vulva is seen frontally, inverted and also from the back, open for penetration. Kurtén sees "unabashed sensual joy and tenderness" in all this. One has the distinct impression that palaeolithic man had a very great interest in and knowledge of sex, and he seems to have portrayed many species including his own in states of sexual arousal. Kurtén must surely be right; it must have been a Golden Age. □

Beverly Halstead is a Reader in the Departments of Geology and Zoology, University of Reading, Whiteknights, Reading RG6 2AH, UK.

• Recently published by the British Museum (Natural History) in the UK, and Facts on File in the United States, is *Mammal Evolution: An Illustrated Guide*, with text by R.J.G. Savage and illustrations by M.R. Long. The book will be reviewed in a future issue of *Nature*.

Cost-effectiveness of short-term tests for carcinogenicity

Lester B. Lave & Gilbert S. Omenn

Most chemicals to which we are exposed are not properly tested for carcinogenicity. The latest methods of in vitro testing provide a way of screening with sufficient accuracy to remedy this situation.

CONCERN about the dangers to health from exposure to toxic chemicals is increasing and all Western nations have developed elaborate schemes for regulating their production and use¹⁻⁶. At present, decisions are made based primarily on long-term animal bioassays¹ which cost more than \$1 million for each chemical and take 2-4 years. Here we present a framework for assessing the effective and efficient regulatory use of inexpensive, short-term *in vitro* tests and illustrate its application with data from an international collaborative testing programme⁷.

The framework

Even though scientific data are rarely conclusive they provide the foundation for regulatory decisions⁸ on whether a chemical should be regarded as carcinogenic or not at a particular exposure level. Carcinogenicity is particularly

difficult to investigate because the causal mechanisms are far from fully understood, the period between exposure and onset of the disease may be decades, data on levels, routes, and time of exposure are generally poor, and susceptibility may vary due to predisposing genetic traits, pre-existing illnesses, personal habits such as cigarette smoking, and exposures to other chemicals that interact with the chemical of interest^{9,10}. Ideally, current data would correctly classify all chemicals into true positives (TP) and true negatives (TN); unfortunately, it is inevitable that some misclassifications, false positives (FP) and false negatives (FN), will occur. Society should attempt to minimize the sum of the costs of testing and of misclassification.

In Fig. 1 we assume that either 2% or 10% of all chemicals are carcinogens. Measuring the cost of additional testing is straightforward, but what are the social costs of misclassification? It is not easy to

generalize about the social costs of continued exposure to a carcinogen (FN). For example, lax standards for asbestos have resulted in hundreds of additional cancer cases each year. In contrast, lowering the benzene standard in the US from 10 to 1 p.p.m. in air was estimated to save less than one cancer per year. Respecting society's deep concern, we put a value of \$10 million on a FN.

Estimating the social cost of a FP may be even harder, since it involves regulating a chemical needlessly. Capital and operating costs of exposure control equipment vary. If chemical production is discontinued, there is loss to the economy and substitutes have to be found. Cyclamates and the detergent nitrilotriacetic acid, both of which have been banned, are probably good examples of FPs. Substitutes may also turn out to have health risks, perhaps even greater ones. We put a cost of \$1 million on a FP. For our purposes, the

a, Assume 2% (1,000) of 50,000 chemicals are carcinogenic

		True biological activity			
		Carcinogenic	Non-carcinogenic		
In vitro test results	Potentially carcinogenic	TP 900	FP 4,900	of false positives 4,900 × \$1 million	= \$4.9
	Not potentially carcinogenic	FN 100	TN 44,100	of false negatives 100 × \$10 million	= \$1.0
				Total	= \$5.9
				Cost of testing	\$0.4
				Total social cost	<u>\$6.3</u>
				Social cost no testing, 1,000 FNs 1,000 × \$10 million	\$10.0

b, Assume 10% (5,000) of 50,000 chemicals are carcinogenic

		True biological activity			
		Carcinogenic	Non-carcinogenic		
In vitro test results	Potentially carcinogenic	TP 4,500	FP 4,500	of false positives 4,500 × \$1 million:	= \$4.5
	Not potentially carcinogenic	FN 500	TN 40,500	of false negatives 500 × \$10 million	= \$5.0
				Total	= \$9.5
				Cost of testing	\$0.4
				Total social cost	<u>\$9.9</u>
				Social cost no testing, 5,000 FNs 5,000 × \$10 million	= \$50.0

Fig. 1 Social costs of testing chemicals for carcinogenicity. Illustrative analysis of consequences of FP (false positive) and FN (false negative) results, assuming that either a, 2% or b, 10% of chemicals are true carcinogens for humans, that the *in vitro* tests have a specificity and sensitivity of 90%, and that the cost of testing is \$8,000 per chemical (test set 6).

critical assumption is that an FN costs ten times as much as an FP.

As shown in Fig. 1, there may be 1,000 or 5,000 unrecognized carcinogens among the 50,000 chemicals in widespread use. The International Agency for Research on Cancer and the US National Toxicology Program list fewer than two dozen recognized human carcinogens^{11,12}. The current cost to society of FNs, therefore, may be \$10,000 million to \$50,000 million (at \$10 million per carcinogen). These figures are disconcerting. They are the hidden costs of an industrial society using chemicals with little knowledge about their potential carcinogenicity (or other adverse effects).

With a test system 90% accurate for both carcinogens and non-carcinogens and costing \$8,000 per chemical, social cost would be \$4,900 million for FP plus \$1,000 million for FN (Fig. 1a). Adding a testing cost of \$400 million, social cost would be \$6,300 million. But if 2% of the 50,000 chemicals in use really are carcinogens, they have a social cost of \$10,000 million. The test system would lower this to \$6,300 million. If 10% are carcinogens, the test system would lower social cost from \$50,000 million to \$9,900 million.

How high could misclassification be and still make the testing worthwhile, under these assumptions? If 2% of chemicals are carcinogens, the classification accuracy must not be lower than 84% for testing to be better than no testing. If 10% are carcinogens, the classification accuracy need only be 52%. Alternatively, if a slightly better test was available, one with 91% accuracy, it would be worth \$11,800 or \$19,000 more per chemical tested, assuming 2% or 10% of chemicals are carcinogens respectively.

The above framework highlights the data required to make better social decisions about toxic chemicals: the cost and accuracy of various carcinogenicity tests; the number of chemicals and proportion that are carcinogenic; and the social costs of false positives and false negatives. Testing 50,000 chemicals in long-term animal bioassays is neither feasible nor cost-effective. It is thus not surprising that increasing effort is being invested in the development and validation of many inexpensive, short-term tests *in vitro*^{7,13-17}. But misunderstandings remain about how accurate these tests would have to be to prove useful for regulatory decision making even though the 'gold standard', the long-term animal bioassay, itself has major limitations, both statistical and biological. To be useful, a single short-term test need not be completely accurate, and the tests need not be useful for all classes of chemicals. Short-term tests are inexpensive to replicate and provide information about dose-response relationships and mechanisms of action of chemicals, rather than simply counting the numbers of tumours in exposed animals.

Table 1 Test chemicals: carcinogen classification based on animal bioassay data⁷

Carcinogens	Non-carcinogens
4-dimethylaminoazobenzene (butter yellow)*	3,3',5,5'-tetramethylbenzidine
chloroform*	4-dimethylaminoazobenzene-4-sulphonic acid Na salt
urethane*	pyrene
DL-ethionine*	α -butyrolactone
hexamethylphosphoramide (HMPA)*	anthracene
ethylenethiourea*	1,1,1-trichloroethane
diethylstilbestrol* (DES)	dimethylformamide
safrole*	diphenylnitrosamine‡
3-aminotriazole*	dinitrosopentamethylene tetramine
<i>o</i> -toluidine hydrochloride*	isopropyl- <i>N</i> -(3-chlorophenyl)carbamate
auramine (technical grade)*	L-methionine
4-nitroquinoline- <i>N</i> -oxide	sucrose
benzidine	L-ascorbic acid
benzo(a)pyrene	4-acetylaminofluorene†
β -propiolactone	3-methyl-4-nitroquinoline- <i>N</i> -oxide†
9,10-dimethylanthracene	1-naphthylamine†
2-naphthylamine	azoxybenzene†
dimethylcarbamoyl chloride	
<i>N</i> -nitrosomorpholine	
methylazoxymethanol acetate	
hydrazine sulphate	
cyclophosphamide	
epichlorohydrine	
4,4'-methylenebis-(2-chloroaniline)(MOCA)	
2-acetyl-amino-fluorene	

* 'Problem carcinogens'.

† Classification changed from non-carcinogen to carcinogen as a result of the testing and excluded from the analysis.

‡ Diphenylnitrosamine was re-classified when animal bioassay data became available, and is now included as a carcinogen.

Test data

The short-term test framework can be illustrated with available test results. The International Program for the Evaluation of Short-Term Tests for Carcinogenicity⁷ was a collaborative study undertaken by the UK Medical Research Council/Health and Safety Executive, the US National Institute of Environmental Health Sciences and Environmental Protection Agency, and the Japanese National Cancer Center Research Institute. We were not involved in the study and gratefully acknowledge the assistance of the program organizers.

Test chemicals were selected to include examples from the major classes of widely used organic compounds. Based on previous animal bioassay results, the 42 chemicals shown in Table 1 were classified as 25 carcinogens and 17 noncarcinogens, including 14 structurally similar carcinogen-noncarcinogen pairs and 11 'problem carcinogens' (reported to be difficult or impossible to detect in standard bacterial mutation assays). The chemicals certainly do not represent a random sample of all chemicals or all widely used chemicals. After changes in classification during testing, a final analysis was based on 38 chemicals; 26 carcinogens and 12 non-carcinogens. Compared with the US Environmental Protection Agency's Genotox data base, this base has proportionally more non-carcinogens¹⁸; it excludes fibres, heavy metals, and some other agents known to cause problems in short- or long-term bioassays.

Thirty assay systems (Table 2) were selected to represent a wide variety of biological endpoints and mechanisms of action but cytogenetic methods may have been under-represented. Some investigators were funded by the study, while other investigators financed their own testing. The investigators were not informed of the identity of the chemicals until after submission of their results. Results were reported as positive, negative, or inconclusive⁷. Unfortunately, quantitative results and data on strength of evidence were generally not reported. Of the chemicals in Table 1, four (*o*-toluidine hydrochloride, HMPA; safrole, DES) are included in the *in vitro* assays and another four (benzo(a)pyrene, pyrene, 2-acetyl-amino-fluorene, 4-acetyl-amino-fluorene) are being subjected to intensive *in vivo* assays in collaborative follow-up studies¹⁸.

Statistical method

We begin by deriving an estimate of the probability that a chemical is carcinogenic from the results of the short-term tests. We define the dependent variable as taking the value of 0 when the chemical is believed to be a noncarcinogen and a value of 1 when it is believed to be a carcinogen. Each independent variable is the report of a test on the chemical; it assumes a value of 0 for negative test results and 1 for positive tests. Inconclusive results and chemicals not tested are excluded from the analysis.

The relationship between dependent variable (presumed carcinogenicity) and

Table 2 Test classification accuracy based on collaborative test results

Test number	Type of test	Cost* (\$)	N†	Accuracy	Test number	Type of test	Cost* (\$)	N†	Accuracy
Bacterial mutation assays					Yeast assays				
1	<i>S. typhimurium</i> /plate	1,200	37	0.70	32	<i>S. cerevisiae</i> XV185-14C	NA	32	0.63
2	<i>S. typhimurium</i> /plate	1,200	36	0.69	33	<i>S. pombe</i> P1	NA	31	0.68
3	<i>S. typhimurium</i> /plate	1,200	34	0.59	34	<i>S. cerevisiae</i> PG-148, PG-154, PG-155, PG-166	NA	29	0.66
4	<i>S. typhimurium</i> /plate	1,200	38	0.71	35	<i>S. cerevisiae</i> D4	NA	35	0.43
5	<i>S. typhimurium</i> /fluctuation	NA	31	0.71	36	<i>S. cerevisiae</i> D6	NA	31	0.42
6	<i>S. typhimurium</i> /plate	1,200	38	0.63	37	<i>S. cerevisiae</i> D7	1,400	35	0.66
7	<i>S. typhimurium</i> /plate	1,200	33	0.64	38	<i>S. cerevisiae</i> JD1	NA	36	0.67
8	<i>S. typhimurium</i> /plate	1,200	38	0.68	39	<i>S. cerevisiae</i> rad	NA	37	0.59
9	<i>S. typhimurium</i> /plate	1,200	37	0.57	In vitro mammalian test systems				
10	<i>S. typhimurium</i> /plate	1,200	37	0.65	40	UDS (unscheduled DNA synthesis) in fibroblasts	5,200	18	0.39
11	<i>S. typhimurium</i> /plate	1,200	37	0.73	41	UDS in fibroblasts	5,200	21	0.62
12	<i>S. typhimurium</i> /plate	1,200	38	0.68	42	UDS in HeLa cells	5,200	38	0.68
13	<i>S. typhimurium</i> /plate	1,200	38	0.71	43	SCE (sister chromatid exchange)	3,000	19	0.53
14	<i>S. typhimurium</i> /+norharman	NA	28	0.57	44	SCE	7,500	18	0.67
15	<i>S. typhimurium</i> /fluctuation	NA	33	0.55	45	SCE	3,000	33	0.52
16	<i>S. typhimurium</i> /azaguanine-res	NA	38	0.66	46	CYT-VTRO	NA	21	0.57
17	<i>S. typhimurium</i> /E. coli WP2/fluc. hepatocytes	NA	36	0.67	47	CYT-VTRO	NA	18	0.72
18	<i>S. typhimurium</i> /E. coli WP2/plate	NA	36	0.78	48	mouse lymphoma L5178Y	4,900	19	0.53
19	E. coli WP2/plate	400	34	0.65	49	CHO-HPRT	NA	9	0.44
20	E. coli 343	NA	38	0.72	50	CHO-HPRT	6,500	3	0.67
Bacterial repair, phage induction, degranulation, nuclear enlargement assays					51	V79-HPRT	6,500	5	0.60
21	<i>B. subtilis</i> , M45 RecA-	800	38	0.79	52	DIPHTH-R	NA	12	0.83
22	<i>E. coli</i> , RecA-/PolA	1,500	34	0.59	53	Transformation	NA	34	0.59
23	<i>E. coli</i> , RecA-	1,500	37	0.62	54	Transformation	NA	38	0.81
24	<i>E. coli</i> , RecA-/PolA	1,500	36	0.58	55	MLV-INFC	NA	24	0.52
25	<i>E. coli</i> , PolA	NA	38	0.59	In vivo assays				
26	λ -induction (<i>gal</i> ⁺)	NA	29	0.65	56	<i>Drosophila</i> -sex-linked recessive lethal	10,000	9	0.44
27	λ -induction (lysis)	NA	36	0.56	57	<i>Drosophila</i> -sex-linked recessive lethal	10,000	15	0.53
28	Degranulation of RER	NA	33	0.39	58	<i>Drosophila</i> -sex-linked recessive lethal	10,000	9	0.44
29	Degranulation of RER + ribonuclease	2,500	30	0.67	59	SCE-in vivo	3,000	17	0.59
30	Nuclear enlargement, HeLa cells	NA	22	0.32	60	Micronucleus	3,400	29	0.66
31	Nuclear enlargement, human fibroblasts	NA	22	0.64	61	Micronucleus	3,400	17	0.47
					62	Micronucleus	3,400	34	0.47
					63	Sperm	11,400	15	0.47
					64	Tradescantia	NA		

Collaborative test result from ref. 7. NA, not available.

* Estimated 1981 costs.

† Number of chemicals tested in the international programme.

independent variables is fitted to the testing results using logit analysis¹⁹. The result is an estimated probability, P , that a chemical is carcinogenic. If the analysis assigned $P = 0$ to all noncarcinogens and $P = 1$ to all carcinogens, the classification would be perfect. With real data, a 'cut-point', P' , is needed to classify all chemicals with $P < P'$ as noncarcinogens and all chemicals with $P > P'$ as carcinogens²⁰. As P' is increased, the number of FPs fall, but FNs appear. The goal is to choose P' to minimize social costs: $kFN + FP$, where k is the cost of an FN relative to that of an FP. If the average social costs of FN and FP were identical ($k = 1$), social cost would be minimized by having the smallest number of misclassifications. However, if FN is valued higher than FP ($k = 10$, as in Figs 1 and 2), then the cut-point is shifted to lower

values (from B to A in Fig. 2). In general P' is significantly greater than zero.

In sum, the statistical procedure has two stages. First, the probability of a chemical being a carcinogen is estimated by using observed short-term test results. Second, a cut-point is sought that minimizes social cost in classifying chemicals.

Reproducibility

Many of the short-term tests employed are relatively new and have not been standardized. In the International Program, twelve investigators used versions of the *Salmonella*/microsome plate assay. Eleven of the 38 chemicals were chosen because previous results in this type of assay system were unclear. No two investigators reported identical results across the 38 chemicals; agreement between investigators ranging from 54% to 97%, even

though the assay is standard. Such discrepancies complicate interpretation and analysis^{21,22}.

Test batteries

Thirty short-term tests were examined in the International Program. They comprise most (but certainly not all) of the short-term tests currently used or proposed for carcinogenicity testing. Notably absent are DNA repair in cultured hepatocytes²³ and various bone marrow cytogenetic tests. Every test was found to be of value in predicting carcinogenicity, with classification accuracies up to 81%, as shown in Table 2. Long-term bioassays may well offer no greater predictive accuracy; among 245 National Cancer Institute bioassays, 55 (22%) were judged inconclusive or equivocal²⁴.

While some tests show impressive

accuracy of classification, a battery of tests should be able to attain even greater accuracy, in large part because a variety of mechanisms can be screened. Several proposals have been made concerning what combination would be best and these are discussed in Tables 2 and 3. Five previous proposals²⁵⁻²⁸ as well as a set we drew from this data base are shown in Table 3 and the results of fitting the data to each test battery are shown in Table 4.

Predicted probabilities of carcinogenicity for each of the 38 chemicals as estimated by test set 6 are shown in Fig. 2. The estimated cut probability for $k=1$ would be between 0.42 and 0.81; for $k=10$, P' would be less than 0.05. The apparent superiority of test set 6 can be ascribed to our method of selecting tests. These results indicate that test 21 (*Bacillus subtilis* Rec-assay) may deserve more attention than it has been given, despite current uncertainties about its mode of action.

Value of testing

How much should society spend to acquire data before making a decision that a particular chemical should be treated as a potential carcinogen in humans? The more nearly complete and the more accurate the data, the lower the chance of an

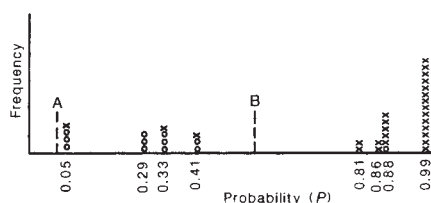


Fig. 2 Predicted probability (P) of carcinogenicity. This histogram for test set 6 shows the frequency of chemicals with P for the set of 38 chemicals in the International Collaborative Program. Cut points corresponding to FN/FP values of 10(A) and 1(B) are indicated by dashed lines. This cut point must be greater than zero to accommodate false positive test results. $\times$ indicates a carcinogen while $\circ$ indicates a noncarcinogen. The three FN (from left to right) are chloroform, auramine, and DL-ethionine; the FP is α -butyrolactone.

FN or FP. The general principle is that data should be gathered until the time and dollar costs of acquiring more data approximate the additional savings to society from lowering the number of FP and FN. Of course, the costs and savings are borne by different parties.

The direct cost of additional data is the price of running more tests. Each test has an associated accuracy and price. Some tests are 'best buys' in the sense that no

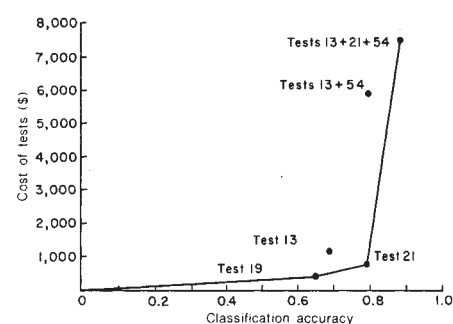


Fig. 3 Cost-effectiveness of *in vitro* testing. The costs of tests are plotted against the estimate of test accuracy listed in Table 1. Tests 13, 21, and 54 correspond to set 6 in Table 2.

test offers the same accuracy for a lower price. Accuracy depends on the sample of chemicals tested. Additional helpful tests give more information, especially when carcinogenic chemicals have heterogeneous mechanisms, but costs rise.

The process of 'buying' additional classification accuracy can be represented schematically in Fig. 3, where the vertical axis represents the cost of testing and the horizontal axis the classification accuracy of the tests in the International Program. In this example, test 13 is both more expensive and less informative than test

Table 3 Test sets analysed

Set number	Test set suggested by	Assays suggested for inclusion	Versions tested
1	Bridges ²⁵	<i>Salmonella</i> /microsome Sex-linked recessive lethal in <i>Drosophila</i> * Malignant transformation Mutation induction in cultured mammalian cells Yeast <i>Neurospora</i> †	<i>Salmonella</i> /microsome (test 13) Cell transformation, BHK-21 (test 54) Forward mutation, <i>S. pombe</i> (test 33)
2	W. L. Marcus, Office of Toxic Substances US EPA ²⁶	<i>Salmonella</i> /microsome <i>S. pombe</i> Mammalian cell culture* Oncogenic transformation Unscheduled DNA synthesis in human cells	<i>Salmonella</i> /microsome (test 13) Forward mutation, <i>S. pombe</i> (test 33) Cell transformation, BHK-21 (test 54)
3	1980 National Toxicology Program Annual Plan ²⁴	<i>Salmonella</i> /microsome Mammalian cell culture Cell culture transformation	Unscheduled DNA synthesis, HeLa (test 42) <i>Salmonella</i> /microsome (test 13) Unscheduled DNA synthesis, HeLa (test 42)
4	I. F. H. Purchase, Imperial Chemical Industries Ltd (see ref. 25)	<i>Salmonella</i> TAI535 & TAI1538 <i>S. cerevisiae</i> * BHK-21 cell transformation Sebaceous-gland suppression† Rabin's test (degranulation)	Cell transformation, BHK-21 (test 54) Rabin's test (degranulation) (test 29)
5	Weisburger and Williams ²⁸	<i>Salmonella</i> /microsome Cell transformations, CHO mammalian mutagenesis Williams' hepatocyte† Sister chromatid exchange in CHO or liver cell systems*	<i>Salmonella</i> /microsome (test 13) Unscheduled DNA synthesis, HeLa (test 42) Cell transformation, BHK-21 (test 54)
6	Our analysis		<i>Salmonella</i> /microsome (test 13) Differential killing, <i>B. subtilis</i> , (test 21) Cell transformation, BHK-21 (test 54)

The short-term tests in Table 2 are grouped into five categories by the mechanisms tested. From the 3 to 13 available assays that represented each category, we constructed a new battery (set 6) by choosing the tests whose results in the International Program data set had explained the highest proportion of variance in the dependent variable (presumptive carcinogenicity). For all batteries except set 3, at least one test had to be omitted.

* Assays included in International Program, but tested on too few chemicals.

† Assay not included in International Program.

Table 4 Logit regression equations

	Test set equation	No. of chemicals	R^2	χ^2	Predictive accuracy	
					C	NC
1	$A = -3.1 + 3.1T_{13} + 2.8T_{33} + 2.8T_{54}$	29	51.9	17.3	4/9	20/20
2	$A = -4.0 + 3.4T_{13} + 2.8T_{33} + 1.8T_{42} + 2.8T_{54}$	29	55.3	19.3	5/9	20/20
3 & 5	$A = -2.8 + 2.3T_{13} + 2.6T_{42} + 3.2T_{54}$	38	48.1	22.6	5/12	26/26
4	$A = -1.8 + 43.9T_{13} + -0.7T_{29} + 2.3T_{54}$	30	44.7	9.6	6/9	20/21
6	$A = -3.1 + 2.2T_{13} + 2.4T_{21} + 2.7T_{54}$	38	56.2	22.9	11/12	23/26

The general logit equation is $P = 1$, where P is the estimated probability that a chemical is a carcinogen and A is a linear equation of estimated coefficients multiplied by test results¹⁷. For example, in set 6, A is calculated as -3.1 plus 2.2 times the result of test 13 (either 0 or 1) plus 2.4 times the result of test 21 (0 or 1) plus 2.7 times the result of test 54 (0 or 1). These coefficients are estimated to give the best fit between the dependent variable, presumptive carcinogenicity, and the P calculated from this logit equation. The coefficient of test 54 (2.7) indicates that this test is more important in predicting carcinogenicity than the other two tests (coefficients 2.2 and 2.4). The omission of some chemicals from the data set and the relatively small size of the sample appear to have led to the nonsense result that test 29 in set 4 should be interpreted as the opposite of the test results. R^2 should be interpreted as the proportion of the variation in the dependent variable accounted for by the regression; the χ^2 statistic shows that the test sets were all statistically significant. Also shown are the proportions of carcinogens (C) and noncarcinogens (NC) (actually tested) that were classified correctly. All the sets perform well, giving classification accuracies up to 89%, although the test batteries selected to be useful in predicting carcinogens are notably less successful in predicting noncarcinogens. Test set 6 correctly predicts 10 of the 11 problem carcinogens known to be difficult for some of these tests to handle.

21. The combination of tests 13 and 54 gives more information at greater cost than 21 alone; the combination of tests 13, 21, and 54 provides more information at only a slight increase in cost. The solid line shows the tests and combinations of tests that provide classification accuracy at lowest cost. For this sample of chemicals, if only about \$1,000 per chemical were available, test 21 would provide 79% classification accuracy. If about \$7,400 per chemical were available, tests 13, 21, and 54 could provide 89% classification accuracy.

Because the classification accuracy of the various test batteries is about 90% and the total cost of testing 50,000 chemicals, with these batteries estimated at \$8,000 per chemical, would be \$400 million, our Fig. 1 is not purely hypothetical. It seems clear that society is making a serious mistake in failing to test most chemicals.

Testing all chemicals immediately is neither feasible nor intelligent. A better approach is a sequential testing policy with orderly selection of priority chemicals for testing (based on production, exposures, and structural similarity to known carcinogens) and use of information on metabolism and mechanisms of action to select tests and interpret results²⁹⁻³². Evaluation of initial test results will lead to modifications of the test regime and classification criteria. Tier-testing systems can address inconclusive or conflicting results. The above discussion is simplistic in dichotomizing chemicals as carcinogenic or harmless, rather than estimating the quantitative toxicity or potency; as testing becomes more sophisticated, it may be possible to estimate potential risks quantitatively.

Conclusion

We have presented a decision analysis framework for the testing and classification of toxic chemicals, using an extraordinary set of published data. The

available set of tests and chemicals is small and difficult to interpret, but it is the best available and illustrates the framework. The set of tests that together predicted the carcinogenicity of the 38 chemicals tested most successfully may not be appropriate for other chemicals, much less the entire population of 50,000 chemicals in use. While the specific results strictly hold only for this limited data set, several general conclusions can be stated.

First, short-term tests can be accurate predictors of the carcinogenicity of chemicals. Sets of tests provided classification accuracy of 89% even when 11 of the 38 chemicals were known to be difficult to detect as mutagenic in the *Salmonella*/microsome assay.

Second, decision analysis modelling of classification problems in terms of the social costs of false positives and false negatives can be instructive. If the social costs are equal, classification accuracy should be maximized. When the social costs of a false negative are greater than those of a false positive, the cut-point should be lowered to classify more chemicals with intermediate probability estimates as carcinogens.

Finally, it may be cheaper for society to test all chemicals in common use with batteries of short-term tests, despite the resulting false positives and false negatives, rather than to continue tolerating thousands of unregulated carcinogens because there are no animal bioassay test results. Several sets of short-term tests appear to be sufficiently accurate and inexpensive to warrant expanded use in testing. However, a more intelligent procedure than routinely testing all 50,000 chemicals would be to test first the highest priority chemicals, based on production and exposures, and to evaluate the test batteries and classification criteria as testing progresses. The US Environmental Protection Agency has authority under the Toxic Substances Control Act to demand

a regime of tests for new chemicals that meet one of two Section 4 criteria: (1) an unreasonable risk of injury to health or the environment based on its physical or biochemical characteristics, or (2) expected production in substantial quantities, giving rise to significant human or environmental exposures.

We recommend that the development and validation of short-term tests for application in screening new and existing chemicals should be accelerated.

Lester B. Lave is James H. Higgins Professor of Economics in the Graduate School of Industrial Administration, Carnegie-Mellon University, Gilbert S. Omenn is Dean of the School of Public Health and Professor of Medicine and Environmental Health, University of Washington. We thank Kathleen D. Heffernan and Glen Dranoff. This work was begun when the authors were at the Brookings Institution. Lave's work was supported by the NSF, Omenn's work by the Sloan and Charles A. Dana Foundations and the National Cancer Institute.

1. Lave, L. B. *The Strategy of Social Regulation* (Brookings, Washington, 1981).
2. Lave, L. B. *Quantitative Risk Assessment for Regulation* (Brookings, Washington, 1982).
3. Calkins, D. R., Dixon, R. L., Gerber, C. B., Zarin, D. & Omenn, G. S. *J. nat. Cancer Inst.* **64**, 169-176 (1980).
4. Ruckelshaus, W. D. *Issues in Science and Technology* **1**, 19-38 (1985).
5. Weinstein, M. C. *Public Policy* **27**, 333-343 (1979).
6. *Regulations Relating to Environmental Chemicals* (OECD, Paris, 1977).
7. deSerres, F. J. & Ashby, J. (eds) *Evaluation of Short-Term Tests for Carcinogens: Report of the International Collaborative Program* (Elsevier, New York, 1981).
8. National Research Council *Risk Analysis in the Federal Government: Managing the Process* (National Academy Press, Washington, 1983).
9. Friedman, R. D. *Sensitive Populations and Environmental Standards: A Legal Analysis* (Conservation Foundation, Washington, 1979).
10. Omenn, G. S. *J. occup. Med.* **24**, 369-374, 1982.
11. IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans: Chemicals, Industrial Processes and Industries Associated with Cancer in Humans (International Agency for Research on Cancer, Lyon, 1982).
12. National Toxicology Program *Third Annual Report on Carcinogens* (US Department of Health and Human Services, Washington, September, 1983).
13. Ames, B. N. *Science* **204**, 587-593 (1979).
14. Office of Science and Technology Policy, *Fed. Register* **49** (100), 21594-21661 (1984).
15. Purchase, I. F. U. *Mutat. Res.* **99**, 53-71 (1982).

16. National Toxicology Program *Review of Current DHHS, DOE, and EPA Research Related to Toxicology* (FY 1983) (US Department of HHS, Washington, 1983).
17. Shelby, M. D. & Stasiewicz, S., *Envir. Mutagen.* **6**, 871-876 (1984).
18. Waters, M. D., Garrett, N. E., Covone-de-Serres, C. M., Howard, B. E. & Stack, H. F. *Genetic Toxicology of Some Known or Suspected Human Carcinogens*, in *Chemical Mutagens: Principles and Methods for their Detection* Vol. 8 (ed. de Serres, F. J.) 261-341 (Plenum, New York, 1983).
19. Goldstein, M. & Dillon, W. R. *Discrete Discriminant Analysis* (Wiley, New York, 1978).
20. Lave, L. B., Omenn, G. S., Heffernan, K. D. & Dranoff, G. *J. Am. Col. Toxicol.* **2**, 125-130 (1983).
21. Chankong, V., Haimes, Y. Y., Rosenkranz, H. S. & Pet-Edwards, J. *Mutat. Res.* **153**, 135-166, 1985.
22. Pet-Edwards, J., Chankong, V., Rosenkranz, H. S. & Haimes, Y. Y. *Mutat. Res.* **153**, 187-200, 1985.
23. Williams, G. D., Laspia, M. F. & Dunkel, V. C. *Mutat. Res.* **97**, 359, 1982.
24. National Toxicology Program *Tech. Bull.* **1**(3), 2 (1980).
25. Bridges, B. S. in *Screening Tests in Chemical Carcinogenesis* (eds Montesano, R., Bartsch, H., Tomatis, L.) 549-568 (International Agency for Research on Cancer, Lyon, 1976).
26. Marcus, W. L. in *Carcinogenesis, A Comprehensive Survey* Vol. 3: *Polynuclear Aromatic Hydrocarbons* (eds Jones, P. W. & Friedenthal, R. I.) 469-472 (Raven, New York, 1978).
27. Office of Technology Assessment *Assessment of Technologies for Determining Cancer Risks from the Environment* 135-136 (US Government Printing Office, Washington, 1981).
28. Weisburger, J. H. & Williams, G. M. *Science* **214**, 401-407 (1981).
29. Hoel, D. G., Kaplan, N. L. & Anderson, M. W. *Science* **219**, 1032-1037 (1983).
30. Crump, K. S., Hoel, D. G., Langley, C. H. & Peto, R. *Cancer Res.* **36**, 2973-2979, 1976.
31. Omenn, G. S. in *Molecular and Cellular Approaches to Understanding Mechanisms of Toxicity* (ed Tashjian, A. H. Jr) 224-245 (Harvard School of Public Health, Boston, 1984).
32. National Research Council *Toxicity Testing* (National Academy Press, Washington, 1984).

ARTICLES

Introduction of homologous DNA sequences into mammalian cells induces mutations in the cognate gene

Kirk R. Thomas & Mario R. Capecchi

Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA

Injection of homologous DNA sequences into nuclei of cultured mammalian cells induces mutations in the cognate chromosomal gene. It appears that these mutations result from incorrect repair of a heteroduplex formed between the introduced and the chromosomal sequence. This phenomenon is termed 'heteroduplex induced mutagenesis'. The high frequency of these events suggests that this method may prove useful for introducing mutations into specific mammalian genes.

GENE-TARGETING (homologous recombination between DNA sequences residing in the chromosome and newly introduced DNA sequences) allows the specific alteration of genes in the mammalian genome¹⁻⁴. Recently, Smithies *et al.*¹ introduced a DNA sequence into the chromosomal β -globin locus by homologous recombination. We have corrected mutant neomycin resistance genes (*neo*) in the host genome via homologous recombination with an injected plasmid DNA carrying a different mutation in the *neo* gene². In the process of analysing these gene-targeting events, we uncovered an unexpected class of 'corrected' chromosomal *neo* genes, which retained the original mutation but had acquired a second mutation that restored gene function. The acquisition of the compensating mutation apparently occurred as a result of incorrect repair of a heteroduplex formed between the *neo* gene in the chromosome and the newly introduced *neo* gene. These reactions occur at a surprisingly high frequency. In addition, we show that the *neo* genes containing the compensating mutations function in mammalian cells and in *Escherichia coli* as a result of a novel translation reinitiation mechanism.

Recombinant plasmids and recipient cell lines

We first established cell lines containing an amber mutant *neo* gene integrated into the genome of the mouse fibroblast line LMtk⁻. We then corrected this mutation by injecting plasmid DNA carrying a non-overlapping deletion mutation in the *neo* gene. Cell lines containing a corrected *neo* gene were identified by selecting for cells resistant to the drug G418.

Figure 1a shows the recombinant plasmids, pRH4-14/TK and pRH140 Δ Nae/TK, used for these experiments. These plasmids were derived from the parental plasmid pRH140³, which contains sequences from the bacterial plasmid pBR322 and the *neo* gene coded by the bacterial Tn5 transposon. The pBR322 sequences supply an ampicillin resistance gene (*amp*^r) and an origin of DNA replication which functions in bacteria. The *neo*

gene was engineered to be functional both in bacteria and in mammalian cells⁵. In bacteria, the *neo*^r gene confers kanamycin resistance; in mammalian cells the *neo*^r gene confers resistance to the drug G418 (G418^r). The herpes simplex virus thymidine kinase gene, HSV-*tk*, was introduced into the above plasmid at the unique *Bam*HI site to generate the plasmid pRH140/TK.

pRH4-14/TK contains an amber codon near the 5' end of the *neo* gene^{5,6}. This premature polypeptide chain termination signal renders the gene product defective in both bacteria and mammalian cells; the mutation also creates a new *Dde*I site which is used as a test for the presence of the amber mutation. pRH140 Δ Nae/TK contains a deletion of 284 bp at the 3' end of the *neo* gene which removes 52 amino acids from the carboxy-terminal end of the NEO protein, rendering it non-functional. The plasmid pRH4-14 Δ Nae/TK contains both the 4-14 amber mutation and the Δ Nae deletion.

We used two recipient mouse cell lines derived from LMtk⁻ fibroblasts. Both cell lines harbour the plasmid pRH4-14/TK, but differ in the copy number and chromosomal location of the integrated plasmid². The cell line LM1 contains a single copy of the plasmid integrated into the host genome by its *Hind*III ends at a locus designated J-1. Cell line LM4 contains four copies of pRH-14/TK integrated at four independent chromosomal sites designated J-1-J-4.

Generation of G418^r cell lines

Plasmid DNA, (~5 molecules per cell) was injected into nuclei of LM1 and LM4 cells, which were then selected with G418. Injection of linear pRH140 Δ Nae/TK molecules into LM1 or LM4 resulted in G418^r cell lines at a frequency of ~1 per 1,000 cells injected (Table 1). This frequency is five orders of magnitude greater than the spontaneous reversion frequency² (to G418^r of either cell line). No G418^r cell lines were obtained following injection of LM1 or LM4 with pBR322/TK, pRH4-

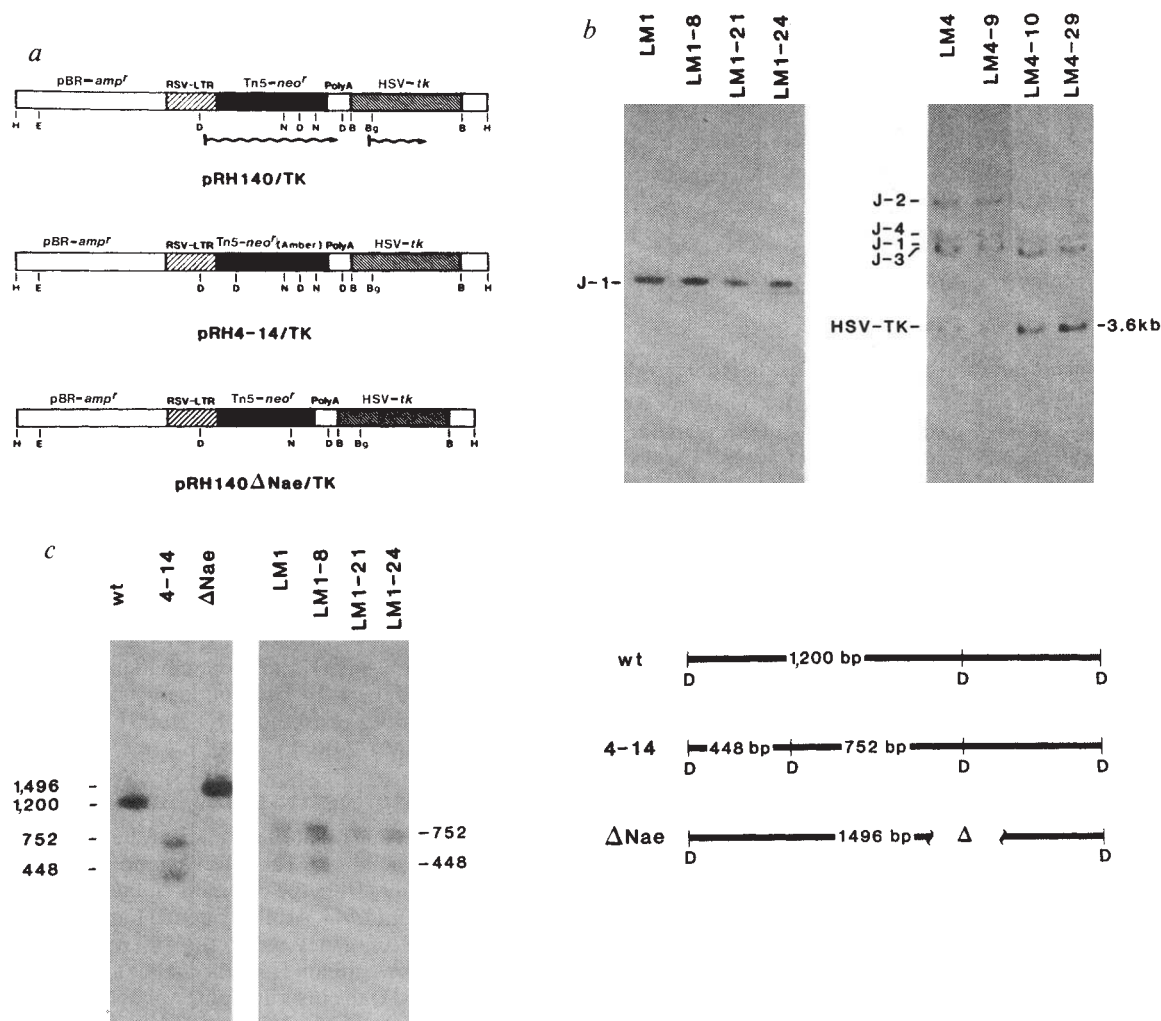


Fig. 1 *a*, Maps of the plasmids pRH140/TK, pRH4-14/TK and pRH140ΔNae/TK. *b*, *Bgl*II and *Bam*HI Southern transfer patterns of LM1, LM4 and their G418^r derivatives. *c*, *Dde*I Southern transfer pattern of LM1, LM1-8, LM1-21 and LM1-24. *a*, The plasmids contain sequences derived from pBR322, the *neo*^r gene coded for by the bacterial Tn5 transposon and the herpes simplex virus thymidine kinase gene (HSV-*tk*). The *neo*^r gene is expressed from a bifunctional promoter, RSV-LTR, which allows expression of the *neo*^r gene in *E. coli* and in mammalian cells⁵. pRH4-14/TK contains an amber mutation that also creates a new *Dde*I site. pRH140ΔNae/TK contains a 284 base pair deletion at the 3' end of the *neo* gene. Each vector is represented in linear form from the unique *Hind*III site. The restriction sites are designated: H, *Hind*III; E, *Eco*RI; Bg, *Bgl*II; N, *Nae*I; B, *Bam*HI and D, *Dde*I. Poly A, polyadenylation sequence from HSV-*tk*. For a more extensive restriction map of these plasmids see refs 2, 6. *b*, DNA was purified from each cell line and digested with either *Bgl*II (LM1, LM1-8, LM1-21, LM1-24) or *Bam*HI (LM4, LM4-9, LM4-10, LM4-29). Aliquots of DNA (5 μg) were electrophoresed through 0.75% agarose, transferred to nitrocellulose paper and probed with either ³²P labelled, nick-translated pRH140 (LM1) or pRH140/TK (LM4). kb, kilobases. *c*, DNA (10 μg) from each cell line was digested with *Dde*I, electrophoresed through 0.75% agarose, transferred to nitrocellulose and probed with a ³²P-labelled, nick-translated, 1200-bp *Dde*I fragment from the *neo*^r gene of pRH140. The sizes (in base pairs) of restriction fragments generated by *Dde*I digestion of the three *neo*^r alleles (wild type, the 4-14 amber mutation and the ΔNae deletion mutation) are given.

14/TK or pRH-14-ΔNae/TK (see Table 1).

Two classes of G418^r cell lines were obtained from LM1 and LM4. In the first class a wild type *neo*^r gene was generated by homologous recombination between the newly introduced pRH140ΔNae/TK plasmid sequence and the pRH4-14/TK chromosomal sequence. This class has been described previously². The second class became resistant to G418^r by virtue of 'heteroduplex induced (het-induced) mutagenesis'.

Fig. 1b shows the Southern transfer patterns of LM1 and LM4 and several of their G418 derivatives. Digestion of genomic DNA with either *Bam*HI or *Bgl*II will isolate the integrated *neo* genes onto single fragments linked to chromosomal sequences. In the case of LM1, *Bgl*II digestion reveals a single hybridizing band representing the single *neo* locus, J-1. The *Bgl*II Southern transfer patterns of the G418^r derivative cell lines LM1-8, LM1-21 and LM1-24, are indistinguishable from the parental pattern. Digestion of LM4 DNA with *Bam*HI reveals

4 bands which hybridize to pRH140 sequences. *Bam*HI digestion of DNA from the G418^r derivatives LM4-9, LM4-10 and LM4-29 show identical patterns. Thus conversion to G418^r was not the result of acquisition of new *neo* sequences or rearrangements of old ones.

As shown in Fig. 1c, digestion of genomic DNA with the restriction enzyme *Dde*I generates a series of fragments that allow the detection of the wild-type *neo*^r gene and each of the *neo* mutant alleles, the 4-14 amber mutation and the ΔNae deletion mutation. The Southern transfer patterns of three G418^r cell lines derived from LM1 following digestion of genomic DNA with *Dde*I are also shown. Only the *Dde*I fragments characteristic of the 4-14 amber mutation are seen. Similar results were obtained when DNA from LM4 and its derivative G418^r cell lines, LM4-9, LM4-10 and LM4-29 were digested with *Dde*I. These results were quite unexpected because all the derivative cell lines were G418^r.

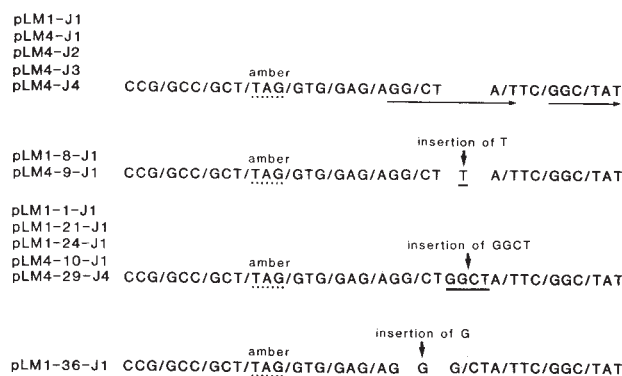


Fig. 2 A summary of pertinent DNA sequences from plasmids rescued from LM1 and LM4 and their G418^r derivatives. Plasmids were rescued from each cell line as described². 5' fragments of the *neo*^r genes from these plasmids were subcloned into the M13 vector, mp8, and sequenced by the chain termination method⁸. The sequences of the protein coding strands representing codons 12–22 are shown. The amber mutations as well as the 6 base pair repeats are underlined.

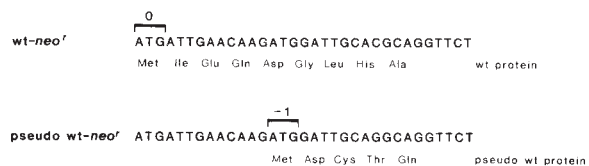


Fig. 3 The amino-terminal sequences of wild-type and pseudo-wild-type *neo*^r proteins are shown with the DNA sequence of the 5' end of the *neo*^r gene. To facilitate protein purification, hybrid proteins were made by in-frame fusions of the 5' ends of *neo*^r genes with the *lacZ* gene. Wild-type fusions were made with the *neo*^r gene from the plasmid pRH140 (ref. 5); pseudo-wild-type fusions were made with the *neo*^r gene from pLM1-1-J1. Both fusions were constructed by inserting a blunt-ended *Bcl*I-*Pvu*II restriction fragment, containing nucleotides -31 to +233 of *neo*^r, into a blunt-ended *Apa*I site in the 8th codon of the *lacZ* gene in p93.94 (R. Weiss, unpublished). *E. coli* cells containing the hybrid genes were grown to stationary phase, lysed in a French Press, and the extract clarified by centrifugation at 280,000g for one h. The β -galactosidase activity was purified either by adhesion to antibodies (mouse anti- β -galactosidase, Promega Biotech), followed by acrylamide gel electrophoresis, or by affinity chromatography⁹. The sequence of the first seven amino acids of each protein was determined on an ABI microprotein sequencer.

Plasmids were rescued from the genome of LM1, LM1-8, LM1-21 and LM1-24. For these experiments genomic DNA was digested with *Bam*HI or *Bgl*II and ligated under conditions that favoured intramolecular ligation. This DNA was used to obtain ampicillin resistant bacteria by transfection⁶. From each cell line we rescued a class of plasmids containing the same 5' junction with chromosomal DNA. Plasmids derived from LM1, as expected, do not provide resistance to kanamycin, but the plasmids rescued from the G418^r cell lines do. When we examined the *Dde*I polymorphism present in these rescued plasmids, all had the 4–14 *Dde*I polymorphism including the plasmids rescued from LM1-8, LM1-21 and LM1-24. Thus all the plasmids retained the amber mutation, yet those rescued from the G418^r cell lines conferred kanamycin resistance on bacteria.

Plasmids were also rescued from LM4, LM4-9, LM4-10 and LM4-29 after digestion of genomic DNA with either *Bam*HI or *Bgl*II. One plasmid from each G418^r line was found to confer kanamycin resistance on *E. coli*. When these plasmids were reintroduced into LMtk⁻ cells G418^r colonies were obtained at a frequency similar to that resulting from injection of the wild-type *neo*^r gene. Thus these plasmids contain a *neo*^r gene that

functions in both bacteria and mammalian cells. The rescued plasmids all had the 4–14 *Dde*I polymorphism, including those that conferred drug resistance (J-1 from LM4-9 and LM4-10 and J-4 from LM4-29).

The compensating mutations are insertions

We sequenced ~300 bp at the 5' ends of the *neo*^r genes isolated from the G418^r cell lines. These were compared with the corresponding sequences from the plasmids isolated from the parental cell lines LM1 and LM4. As illustrated in Fig. 2, the only changes found in the *neo*^r genes were: the insertion of a thymidine residue 11 bp downstream from the amber mutation in pLM1-8-J1 and pLM4-9-J1; the insertion of four bases, GGCT, in pLM1-1-J1, pLM1-21-J1, pLM1-24-J1, pLM4-10-J1 and pLM4-29-J4; and the insertion of a guanosine residue 8 or 9 bp downstream from the amber mutation in pLM1-36-J1.

Reinitiation of translation

To determine how *neo*^r genes containing both an amber mutation and a frameshift mutation confer *neo*^r activity in both bacteria and mammalian cells, we analysed the amino-terminal sequence of the *neo*^r gene products. We will refer to the gene containing both the amber mutation and an insertion mutation as the pseudo-wild-type gene. To facilitate analysis, the 5' 200 nucleotides of the *neo* genes were fused to the *lacZ* gene. Hybrid β -galactosidase containing the amino termini from the wild-type and pseudo-wild-type NEO protein containing the amber mutation and the GGCT insertion were purified from transformed *E. coli* and subjected to amino-terminal sequencing. The fusion protein from wild-type *neo*^r-*lacZ* gene has the sequence Met-Ile-Glu-Gln-Asp-Gly... as predicted (Fig. 3). However, the protein from the pseudo-wild-type *lacZ* gene fusion began with the sequence Met-Asp-Cys-Thr-Gln. This protein was initiated from an AUG codon 14 bp downstream from the normal AUG codon in the -1 translation reading frame (see Fig. 3). It now becomes clear how the insertion mutations reverse the amber mutation. After initiation of protein synthesis in the -1 frame, the ribosome passes through the amber mutation, which is in the 0 reading frame and therefore not read, and regains proper phase when it reaches the +1 frameshift mutation (insertion of a T, G or GGCT). Although this alters the amino-terminal amino acid sequence of the NEO protein, this portion of the protein has been shown to be dispensable¹⁰.

Figure 4a shows two models that could explain how translation of the pseudo-wild-type gene begins at the -1 AUG codon.

Table 1 Transformation frequencies to G418^r by injecting LM1 and LM4 cells with plasmid DNAs

DNA injected	Frequency of G418 ^r cell lines generated by hetero-duplex induced mutagenesis	Number of cells injected
pBR/TK	0	2 × 10 ⁴
pRH140ΔNae/TK	1.2 × 10 ⁻³	10 ⁴
pRH4-14/TK	0	3 × 10 ⁴
pRH4-14-ΔNae/TK	0	2 × 10 ⁴

LM1 or LM4 cells were grown on glass cover slips (10 mm × 10 mm) in 35 mm Petri dishes. Twenty-five cells per dish received nuclear injections of ~5 plasmid molecules per cell⁷. Plasmid molecules were linearized with *Hind*III (pBR/TK) or *Bcl*I (pRH140ΔNae/TK; pRH-14/TK; pRH4-14-ΔNae/TK). After the injection, the cells were incubated for 24 h in nonselective medium at 37° in a 5% CO₂ incubator and then switched to minimum essential medium supplemented with 400 μg ml⁻¹ G418. The dishes were scored for colonies after 3 weeks. pBR/TK contains the 3.6 kb *Bam*HI fragment of the herpes *tk* gene inserted at the *Bam*HI site of pBR322. pRH4-14/TK and pRH140ΔNae/TK are described in Fig. 1 legend. pRH4-14-ΔNae/TK was created by the deletion of the 284 bp *Nae* fragment from the *neo*^r gene in pRH4-14/TK.

In model I we assume that the ribosome can enter at two sites, AUG codons in the 0 and -1 reading frame. Model II uses a single entry point followed by translation reinitiation. In this model the ribosome enters at the AUG codon in the 0 reading frame and translates the messenger RNA until it reaches the amber codon. There it terminates and releases the polypeptide fragment. The ribosome then scans back until it reaches the -1 AUG codon, where it reinitiates protein synthesis.

We performed two sets of experiments to distinguish between these models. First, the amber mutation was removed, by *in vitro* site-directed mutagenesis, from the pseudo-wild-type gene. Second, a four bp insertion mutation was created *in vitro* in the wild-type *neo^r* gene just downstream from the -1 AUG codon. Plasmids containing these altered pseudo-wild-type and wild-type genes were introduced into *E. coli* and cultured mammalian cells to evaluate their ability to confer kanamycin and G418 resistance (see Fig. 4b). The two models make opposite predictions of the resultant phenotypes. If the ribosome can enter at either AUG then both the above plasmids should be functional; the ribosomes entering at the -1 AUG codon would regain proper phase on reaching either of the +1 frameshift mutations. However, neither set of plasmids would be functional if only the initial entry site is used because: (1) after entering the pseudo-wild-type mRNA lacking the amber mutation, the ribosome would be unable to terminate and therefore unable to reinitiate protein synthesis; (2) after entering a wild type mRNA containing a four base-pair insertion, the ribosome would encounter a +1 frameshift and generate a mutant gene product. As shown in Fig. 4b, neither of these *neo* genes is functional in either bacteria or mammalian cells. These results strongly support the single entry-termination-reinitiation model.

Discussion

The unexpected finding that correction of the pRH4-14/TK sequences in the host genome of LM1 and LM4 often results from insertion of a few base pairs downstream from the amber mutation raises a number of questions. How did the insertions of these base pairs occur? How did injection of pRH140ΔNae/TK into either LM1 or LM4 mediate these events? The DNA sequence surrounding the position of the base-pair insertions contains the six bp direct repeat GGCTAT (see Fig. 2). The fact that each of the insertions, T, G or GGCT is a tandem repeat of part of this sequence suggests a duplication-generating mechanism.

A number of observations argue that the insertion or duplication events at this site were induced by the initiation of recombination between the pRH4-14/TK sequence residing in the chromosome and the incoming pRH140ΔNae/TK sequence. First, the frequency of generating this class of G418^r cell lines was comparable to the frequency of generating cell lines by legitimate gene replacement or gene conversion². Second, this frequency is five orders of magnitude greater than the spontaneous reversion frequency of LM1 or LM4 to G418^r. Third, we previously isolated a cell line in which both a homologous recombination event and a het-induced mutagenic event occurred². Because the homologous recombination events and the mutagenic events each occur at a frequency of ~1 per 1000 cells injected, the predicted frequency of both events occurring independently in the same cell line is 1 per 10⁶ injected cells. As we obtained such a cell line after a few thousand injections, it is tempting to postulate that the two events occurred as a result of a concerted reaction. Fourth, we did not obtain G418^r cells from LM1 or LM4 following injection of a plasmid DNA, pBR322/TK, that does not contain sequences similar to the *neo* gene, so injection of DNA does not *per se* induce rampant mutagenesis. Similarly, we did not obtain G418^r cells from LM1 or LM4 following injection of pRH4-14/TK or pRH4-14ΔNae/TK. These experiments demonstrate the specificity of the reaction. Because injection of the pRH4-14/TK vector does not induce the mutations, an important factor for triggering the

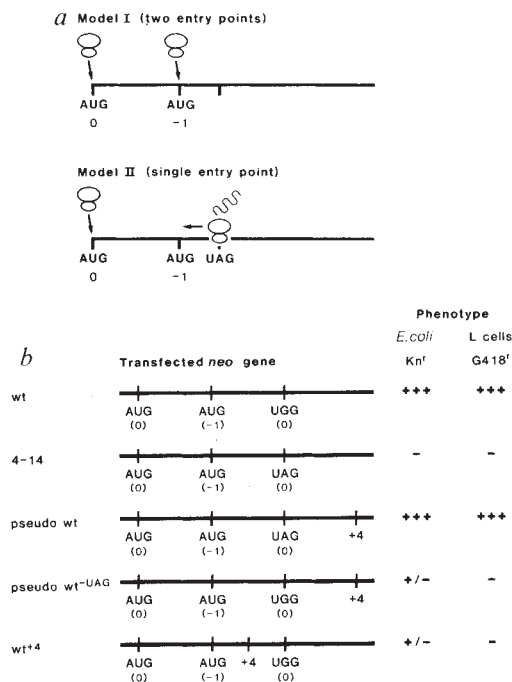


Fig. 4 *a* Two models explaining the synthesis of pseudo-wild-type protein. *b*, Phenotypic analysis of transfected *neo^r* genes. *a*, In model I, ribosomes entering *neo^r* messenger RNA bind at two binding sites, one on the 0 reading frame at nucleotide 1, the other on the -1 reading frame at nucleotide 14. Proteins made from the latter site would be shifted onto the appropriate reading frame by downstream +1 frameshifts. Model II proposes a single ribosome entry site in the 0 reading frame at nucleotide 1. To use the -1 AUG codon at position 14, the ribosome must first translate in the 0 reading frame up to the UAG (amber) codon. At this point, the ribosome will release the newly synthesized peptide and then scan the message until it locates the alternative AUG codon in the -1 reading frame. *b*, Plasmids containing various *neo^r* genes were assayed for the ability to confer drug resistance to *E. coli* or to mouse L cells. The wild type *neo^r* gene (wt) from pRH140 contains a tryptophan codon, UGG, at codon 15. 4-14 is the amber mutation from the plasmid pRH4-14 and contains UAG at codon 15. Pseudo-wt is a sequence containing both the 4-14 amber mutation and a downstream, four base-pair insertion of GGCT, following nucleotide 56. Pseudo-wt^(-UAG) was derived from the pseudo-wt gene, by removal of the UAG codon by site-directed mutagenesis¹¹. The mutant wt⁺⁴ was derived from the wild-type gene following the insertion of four base pairs at position 35. GGCC was inserted at this position by filling in the ends of an *Xma*III restriction site. The figure depicts the 5' end of the *neo^r* mRNA from each plasmid. AUG (0), the translational initiation codon in the 0 reading frame; AUG (-1), the proposed initiation codon in the -1 reading frame at nucleotide 14; the sites of insertional mutagenesis of GGCC following base 35 or GGCT following base 56 are indicated (+4); UGG and UAG are the alternative states of the fifteenth codon. Plasmids containing these mutations were introduced into *E. coli* strain MH1 by CaCl₂-mediated transformation or into mouse L cells by nuclear microinjection. Sensitivity of *E. coli* to kanamycin or of L cells to G418, was determined as previously described⁶. +/-. Resistance to low levels of kanamycin. We believe that this low level of resistance was again due to reinitiation of the -1 AUG as these +1 frameshift mutations generate a new nonsense mutation in the +1 reading frame much further downstream from the -1 AUG (a UGA codon 53 bp downstream from the -1 AUG codon).

mutagenic response may be the single base-pair mismatches at the amber mutation and/or the large mismatch at the ΔNae deletion. The observation that injection of the double mutant, pRH4-14-ΔNae/TK, also does not induce mutations argues that at least the single base-pair mismatch is required. The only mechanism that we can envision by which single base-pair mismatches between a chromosomal sequence and an

exogenous sequence can influence the mutation process requires formation of a heteroduplex.

A number of models can be drawn in which the strands in the heteroduplex transiently misalign at the direct repeats leading to partial duplication of this sequence as a consequence of DNA repair. A model mechanism for the insertions found in the LM1 and LM4 G418^r cell lines is given in Fig. 5. A misaligned heteroduplex between the *neo* gene in the chromosome and the injected *neo* gene is formed; a nick is then introduced near one of the loops, generating a primer for strand extension. The condition that the insertion mutation produces a functional gene restricts both the position of the nicks and the number of nucleotides that can be incorporated at the nick. We have shown that insertions that generate a frameshift mutation in the +1 translation reading frame will compensate for the upstream amber mutation. This condition limits the number of bases incorporated at the nick to one base, 4 bases, 7 bases etc. The position of the nicks is also restricted. Nicks at some positions followed by insertion of one or four bp introduce new nonsense mutations. Thus, only a limited set of possible mutations exist which generate a functional *neo*^r gene. After primer extension, the DNA strands are ligated and the insertion fixed into the genome by a round of DNA replication. In prokaryotic recombination and/or DNA repair, reactions homologous to the ligation depicted in the model may not occur; however, such reactions do occur in mammalian cells¹².

We stress that all the insertion mutations that we analysed resulted from independent events. They involve insertions of different bases, in different cell lines, at different chromosomal loci and were generated at different times. It is interesting that all eight insertions occurred in the first repeat. The proximity and/or the nature of the mismatch at the amber mutation may account for this polarity.

The translation reinitiation mechanism described in this paper could be a general mechanism for translating polycistronic messenger RNAs in eukaryotic cells. After the termination of the first protein, the ribosome could scan, forward or backwards, for the AUG codon used to initiate translation of the second protein. Recent reports^{13,14} support such a model for translation of a polycistronic messenger. Further, the *src* mRNA of the Rous sarcoma virus and some genes of the cauliflower mosaic virus may be translated by such a mechanism^{15,16}. A similar mechanism can also explain the observation^{17,18} made with *in vitro* engineered templates, that translation-initiation at an internal AUG codon is stimulated by the presence of a termination codon in frame with the upstream AUG codon.

Conclusion

As a consequence of correcting a gene residing in the host genome by 'gene-targeting', we uncovered an unexpected class of corrected genes. These genes still retained the original mutation but acquired a compensating mutation. The frequency of this event was surprising and was shown to depend not only on homology but also on mismatched base pairs between the newly introduced DNA sequence and the corresponding sequence residing in the genome. All the insertion mutations occurred in a small direct repeat proximal to the base-pair mismatch, suggesting that the direct repeat is also an important component for generating high frequency 'het-induced mutagenesis'. The actual frequency of mutagenesis may be much higher than the observed frequency of 1 per 1000 cells receiving an injection because we imposed the condition that only mutations that

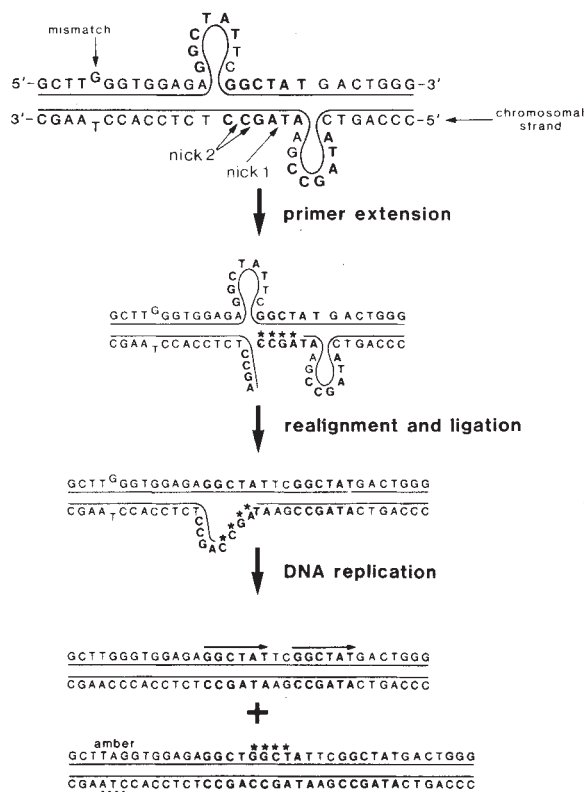


Fig. 5 A model for the mechanism of inserting GGCT into the chromosomal pRH-14/TK sequence. The first step features the formation of a misaligned heteroduplex at the six base-pair direct repeat, between the *neo*^r gene in the chromosome and the injected *neo*^r gene. In the second step a nick is introduced near one of the displaced loops to generate a primer for strand extension. Following extension, the strands are ligated and the insertion fixed into the genome by a round of DNA replication. The positions of the nicks were chosen to account for the insertions isolated: nick 1 would generate either the +T or the +GGCT insertions; nick 2 would generate the +G insertion. The site of the nick and the length of strand extension is limited by the requirement to produce a functional gene. Bold letters, the bases involved in the six base-pair repeats; *, the bases added during strand extension.

converted the amber mutant into a functional gene will be identified. Most such mutations would not be expected to restore gene activity.

In future, we will examine 'het-induced mutagenesis' that directs the loss of gene function rather than the correction of gene function. These events should occur at a higher frequency and reveal a wider spectrum of changes at the DNA sequence level. The only apparent requirements for a 'hot spot' for het-induced mutagenesis are small direct repeats proximal to base-pair mismatches between the newly introduced plasmid sequence and the homologous sequence in the genome. Comparable small direct repeats are encountered in the DNA coding sequence of most genes, making them susceptible to this type of mutagenesis. Permutations of this methodology should provide the means for efficiently introducing mutations into specific mammalian cellular genes.

We thank Laurie Fraser for assistance with tissue culture and Regina Zeikus, Robert Weiss and Diane Dunn for assistance with protein sequencing.

Received 18 July; accepted 2 October 1986.

- Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. S. *Nature* **317**, 230-234 (1985).
- Thomas, K. R., Folger, K. R. & Capecchi, M. R. *Cell* **44**, 419-428 (1986).
- Lin, F. L., Sperle, K. and Sternberg, N. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1391-1395 (1985).
- Smith, A. J. H. & Berg, P. *Cold Spring Harb. Symp. Quant. Biol.* **49**, 171-181 (1984).
- Hudziak, R. *et al. Cell* **31**, 137-146 (1982).
- Folger, K. R., Thomas, K. R. & Capecchi, M. R. *Molec. cell. Biol.* **5**, 59-69 (1985).
- Capecchi, M. R. *Cell* **22**, 479-488 (1980).
- Sanger, F., Nicklen, S. & Coulson, R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).

- Ullman, A. *Gene* **29**, 27-31 (1984).
- Beck, E., Ludwig, G., Auerswald, E. A., Riess, B. & Schaller, H. *Gene* **19**, 327-336 (1982).
- Hutchinson, C. A. *et al. J. biol. Chem.* **253**, 6551-6560 (1978).
- Roth, D. B., Porter, T. N. & Wilson, J. H. *Molec. cell. Biol.* **5**, 2599-2607 (1985).
- Peabody, D. S. & Berg, P. *Molec. cell. Biol.* **6**, 2695-2703 (1986).
- Peabody, D. S., Subramani, S. & Berg, P. *Molec. cell. Biol.* **6**, 2704-2711 (1986).
- Hughes, S. *et al. Molec. cell. Biol.* **4**, 1738-1746 (1984).
- Dixon, L. K. & Hohn, T. *EMBO J.* **3**, 2731-2736 (1984).
- Johansen, H., Schumperli, D. & Rosenberg, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7698 (1984).
- Liu, C., Simonsen, C. C. & Levinson, A. D. *Nature* **309**, 82-85 (1984).

Outbursts of Cygnus X-3 observed at 1.3 and 3.3 mm wavelengths

J. W. M. Baars*, W. J. Altenhoff*, H. Hein†
& H. Steppe†

* Max-Planck-Institut für Radioastronomie, D-5300 Bonn 1, FRG

† Instituto de Radioastronomía Millimétrica, Avén. divina Pastora 7, 18012, Granada, Spain

The X-ray source Cyg X-3 is associated with a radio source which shows irregular, fast and extremely strong outbursts of radiation. The first such event was seen in September 1972 by many observers over a wide frequency range¹; several more events have occurred since then. The main characteristics of the outbursts have been described by the behaviour of an expanding synchrotron source, which undergoes adiabatic and synchrotron losses^{1,8}. However, the detailed physics of the outbursts remains to be understood, in spite of a growing set of observational data. Here we report observations of Cyg X-3 made simultaneously at wavelengths of 1.3 and 3.3 mm, showing outbursts with an intensity of several Janskys and a duration of only a few hours.

Cyg X-3 is a binary object with a 4.8-h periodicity in both X-ray and infrared emission¹⁰. This periodicity may also be visible in low-level radio flares², but was not observed in several outbursts at a level above 50 mJy (ref. 3). Following a suggestion by K. Johnston that a flare was imminent, we included Cyg X-3 as a test object for continuum observations, during the commissioning of the new 30-m-diameter Millimetre Radio Telescope⁴ of the French-German Institute for Radioastronomy in the Millimetre Range (IRAM).

Apart from an observation during the September 1972 outburst at 3.3 mm (ref. 5), no data are available on the development of a burst at millimetre wavelengths. Our observations made simultaneously at two wavelengths (1.3 and 3.3 mm) enable us to study the time dependence of the spectral index. We observed Cyg X-3 from 1 to 11 October 1985. On 1/2 October we barely detected Cyg X-3, whereas on 5/6 October it was easily seen. Subsequently, we followed the object on two consecutive days, from 17:00 UT on 6 October until 02:00 UT on 7 October, and from 16:00 UT on 7 October until 02:45 UT on 8 October. On 11 October a single observation at 3.3 mm was obtained.

The IRAM 30-m telescope is located at 2,850 m altitude on Pico Veleta in the Sierra Nevada near Granada, Spain. The telescope has an altazimuth mounting with a Nasmyth focus in a spacious instrument cabin. During our observations, the beams of the two receivers were coincident to within 2 arc s on the sky.

The receivers were equipped with cooled Schottky-mixers and had a double-sideband system temperature (excluding noise from the antenna and the atmosphere) of 360 K and 200 K at the wavelengths of 1.3 and 3.3 mm (frequencies of 226 and 90 GHz), respectively, and a bandwidth of 500 MHz. Calibration of the system sensitivity is achieved with the aid of loads at ambient and liquid-nitrogen temperatures, introduced in the plane between the quaternary Nasmyth mirror and the receiver input.

At millimetre wavelengths the sensitivity is usually limited by fluctuations in the lower atmosphere. These can be effectively suppressed by quickly moving the antenna beam on and off the source and measuring the difference in the signal received from the two directions. In the lower region of the atmosphere, the two beams have an almost perfect overlap, so that the fluctuations are nearly cancelled in the difference signal⁶.

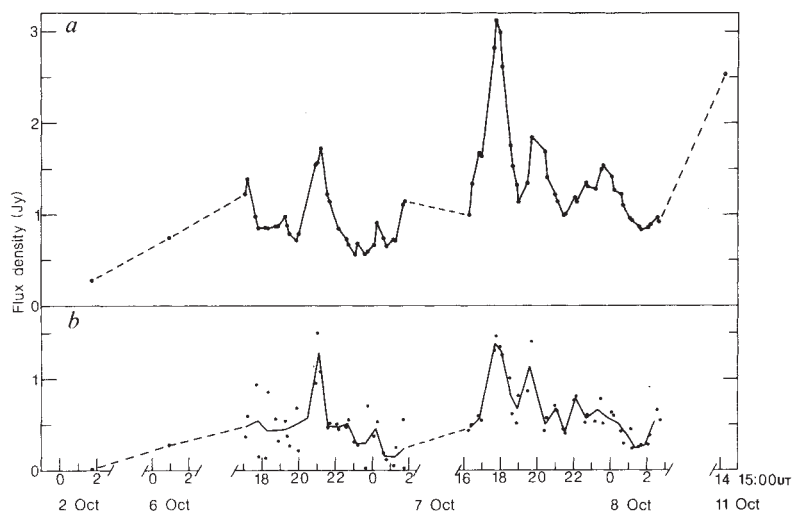
For these 'beam-switched' observations we installed a focal-plane chopper. A vane in the form of a butterfly with a diameter of 1.5 m, rotated parallel to and in front of the quaternary mirror, moved the beam through an angle of 82 arc s on the sky at a rate of 6 Hz. Excellent suppression of atmospheric fluctuations was achieved. The half-power width of the antenna beam is 26 arc s at 3.3 mm wavelength and 12.3 arc s at 1.3 mm, and during our observing period, the aperture efficiencies at these wavelengths were 0.53 and 0.24, respectively.

Our method of observation consisted of measuring the ratio of the intensities of Cyg X-3 and the planetary nebula NGC7027. The angular distance between these two sources is only a few degrees, so that any dependence on the atmospheric extinction and antenna efficiency is eliminated from the ratio. NGC7027 is strong enough to be used for measuring pointing corrections, and, moreover, does not show any time variations in its flux density. We used Mars as a flux density calibrator. On 9 November 1985 we determined the ratio NGC7027/Mars, when Mars was near aphelion and had an apparent diameter of 4.04 arc s. We assumed a black-body brightness temperature of 198 K with an absolute error of probably <5% (ref. 7), and derived flux densities of 93.8 and 14.9 Jy, leading to observed peak flux densities for NGC7027 of 4.27 and 4.80 Jy at 1.3 and 3.3 mm, respectively.

Following a pointing check on NGC7027, its intensity was determined by a set of five on-off pairs, in which the on- and off-axis beams were alternately placed on the source for an integration time of 10 s each. Then a similar set of ten on-off pairs was obtained for Cyg X-3. Such a cycle lasted for 30 min, after which it was repeated.

Figure 1 shows the derived flux densities of Cyg X-3 as a function of time. The noise in a set of ten on-off measurements (200 s of integration time) is ~0.1–0.2 and 0.04 Jy at 1.3 and

Fig. 1 Flux density of Cyg X-3 during 2–11 October 1985 at wavelengths of 3.3 (a) and 1.3 (b) mm, observed with the IRAM 30-m telescope. In b the line connects the averages of two adjacent measurements, which are shown individually as dots.



3.3 mm, respectively. By combining pairs of data points the 1.3 mm results were smoothed considerably.

During our observations three strong outbursts occurred almost simultaneously at both wavelengths. The first and last are nearly equally strong at both frequencies (spectral index $\alpha = -0.2$, where $S \propto \nu^\alpha$ for flux density S and frequency ν), whereas the middle outburst is twice as strong at 3.3 mm ($\alpha = -0.7$). The average plateau-level flux density also has $\alpha = -0.7$ on both days.

During the observation period the flux density of Cyg X-3 was monitored at 11 cm wavelength with the Green Bank interferometer (K. J. Johnston, personal communication). Our first observation on 1/2 October coincides with the beginning of the precursor burst at 11 cm, and the main flare at 11 cm rose from 2.0 Jy at 00:00 UT on 6 October to 9.0 Jy at 00:00 UT on 8 October. Our millimetre-wavelength observations show only a slow increase of the plateau level during this interval.

The time development of the 11 cm flare is similar to earlier outbursts observed at wavelengths between 2 and 21 cm. The first widely observed flare of September 1972 has been well described by a model of an expanding cloud of relativistic particles, mixed with an ionized thermal gas⁸, in which the flare reaches its maximum earlier at shorter wavelengths.

This model is clearly not appropriate for the character of the flare at millimetre wavelengths. The duration of the millimetre-wavelength bursts is only ~ 1 h between the half-power points, with a suggestion of a broadening at 3.3 mm. This could explain why these single, short outbursts are not seen at the longer centimetre wavelengths.

The millimetre-wavelength bursts are probably related to a repeated injection of relativistic particles. Such a scheme is necessary to adequately describe the early phase of the centimetre-wavelength flares⁹. At injection the electrons will have the highest energy, and in losing their energy during the expansion, they will radiate first at the shortest wavelengths. The millimetre-wavelength flares of short duration may be manifestations of separate bursts of injected particles, in which case differences in the optical depth of the medium could cause the observed differences in spectral index. To understand how these short bursts build up to the observed centimetre-wavelength flare will require the analysis of detailed, simultaneous observations at millimetre and short centimetre wavelengths.

We have looked in these data for the 4.8-h periodicity present in the X-ray and infrared emission of Cyg X-3. A spectral analysis indicates a possible periodicity of ~ 5 h in the 3.3-mm data on 6/7 October, which can be seen also in Fig. 1. There is a suggestion of a 2.5-h periodicity in the 1.3-mm data on 7/8 October. Because our observations on 7 October started almost exactly 3×4.8 h after we stopped observing on 6/7 October, we also combined the data and searched the total 20 h of data for a 4.8-h periodicity without finding a significant result. We conclude that our data do not produce evidence for a 4.8-h periodicity in the millimetre-wavelength outbursts on 6–8 October 1985, similar to the result in ref. 3 for centimetre-wavelength outbursts above 50 mJy. Our noise level is too high to detect a possible periodicity in the low-level radiation².

We thank K. J. Johnston for a timely warning of a possible outburst and for data before publication. The observation on 11 October was made by A. Baudry.

Received 6 May; accepted 31 July 1986.

1. Gregory, P. C. *et al.* *Nature* **239**, 440–443 (1972).
2. Molnar, L. A., Reid, M. J. & Grindlay, J. E. *Nature* **310**, 662–665 (1984).
3. Johnston, K. J. *et al.* *Astrophys. J.* (submitted).
4. Baars, J. W. M., Hooghoudt, B. G., Mezger, P. G. & de Jonge, M. J. *Astr. Astrophys.* (submitted).
5. Pomphrey, R. B., Epstein, E. E. *Nature phys. Sci.* **239**, 125–126 (1972).
6. Baars, J. W. M. *Nature* **212**, 494–495 (1966).
7. Ulich, B. L. *Astr. J.* **86**, 1619–1626 (1981).
8. Marscher, A. P. & Brown, R. L. *Astrophys. J.* **200**, 719–726 (1975).
9. Peterson, F. W. *Nature* **242**, 173–177 (1973).
10. Parsignault, D. R. *et al.* *Nature Phys. Sci.* **239**, 123–125 (1972).

Io plasma torus and the source of jovian kilometric radiation (bKOM)

Dyfrig Jones

Space Plasma Physics Group, British Antarctic Survey, NERC, Madingley Road, Cambridge CB3 0ET, UK

One of the many discoveries by Voyagers 1 and 2 during their flyby of Jupiter was the detection of radio emission at kilometric wavelengths^{1,2} in the frequency range 20–500 kHz. Broadband jovian kilometric radiation bKOM was observed to be beamed away from the zenomagnetic equatorial plane, the beaming angle increasing with increasing frequency^{1–3}. It was suggested that bKOM is created by the linear conversion of electrostatic upper hybrid waves to electromagnetic waves in the density gradients at the Io torus⁴. The beaming is inherent to the conversion mechanism and depends on the characteristic frequencies of the source plasma. Here I present results, based on this theory, of remote sensing of the Io torus using bKOM observed by Voyager 2 when inbound towards Jupiter. The density profiles obtained are in remarkable agreement with measurements of the torus densities obtained *in situ* by Voyager 1. This seems to provide very strong evidence that bKOM is produced at the Io torus by the linear mode-conversion mechanism and indicates that the technique used for remote sensing could be a powerful method of determining the torus characteristics when no *in situ* measurements are available.

bKOM first appeared in the planetary radio astronomy (PRA) data when Voyagers were >2 AU from the planet^{1,2}. Figure 1 shows a PRA wave spectrogram⁵ in which bKOM is the tapered emission recorded when Jupiter's north magnetic pole (located at central meridian longitude, CML = 202°) is tilted towards the spacecraft. The leading edges of the tapered emissions occur when Voyager 2 is at CML $< 202^\circ$ and the trailing edges when Voyager's CML is $> 202^\circ$. Voyager 2's jovian latitude when inbound (at 09:30 h jovian local time, LT), was 7.2° and, because Jupiter's magnetic dipole axis is tilted at 9.6° to its rotation axis, the spacecraft's excursion in magnetic latitude during one jovian rotation was -2.4° to 16.8° . Spectrograms such as that in Fig. 1 show that bKOM was generally observed when Voyager was at magnetic latitudes $> 10^\circ$ N. Voyager 1, whose magnetic latitude varied between -6.5° and 12.7° when inbound (at 10:30 LT), and which therefore went to more southerly latitudes than Voyager 2, occasionally observed a similar beaming in the southern hemisphere⁶. This would indicate approximate symmetry about the magnetic equator with a wedge-shaped region of no illumination near the equatorial plane, the width of the region depending on frequency. Contrary to earlier reports⁵ the radiation is right-hand polarized in the northern hemisphere and left-hand polarized in the south⁶. The tapered form of the emission evident in Fig. 1 clearly shows that the higher frequencies are beamed to higher latitudes.

The linear conversion theory⁴ proposed for the production of bKOM invoked electrostatic upper hybrid (ESUH) waves as the source, with their propagation in the Io plasma torus ultimately resulting in the emission of the electromagnetic O mode radiation. Intense ESUH waves were observed within $23 R_J$ of the planet which were confined to within 1.9° of the magnetic equatorial plane⁷. In the theory, it is proposed that these waves propagate in the torus density gradient which is assumed to be normal to the magnetic field. The ESUH waves first become electromagnetic Z mode waves⁸ which, in turn, after further propagation, convert into electromagnetic O mode radiation through a radio window⁹ which exists where the wave frequency equals the plasma frequency $f_p = 9 \times 10^{-3} N_e^{1/2}$ kHz, $N_e \text{ m}^{-3}$ being the electron density. On passing through the radio window into the low-density region outside the torus the O mode radiation emanates as two beams at angle $\beta = \arctan(f_p/f_c)^{1/2}$ to

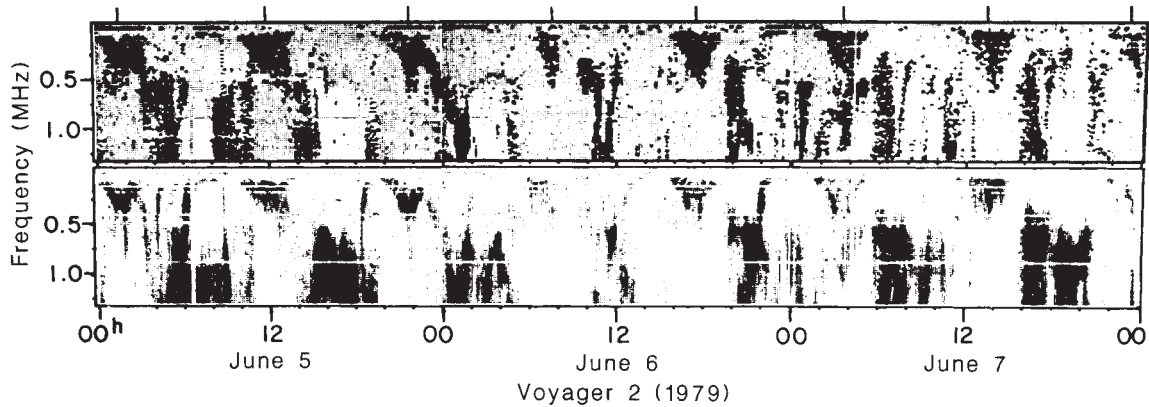


Fig. 1 Voyager 2 spectrogram⁵ of jovian radio emissions in the range 20 kHz–1.3 MHz. Total power is shown in bottom panel, increasing darkness being proportional to increasing intensity. Upper panel shows polarization information on the same emissions. In the 72-h period shown, eight consecutive bKOM events are visible as the V shaped emissions extending up to ~500–600 kHz and coincident with the north dipole tip passage: their polarization is right-handed⁶. The events are anticoincident with the hectometric radiation seen at higher frequencies.

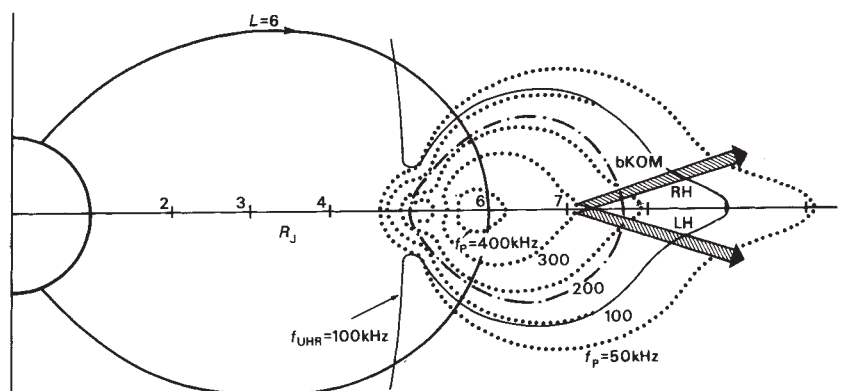
the magnetic field line⁴, $f_c = 0.028 B$ kHz being the electron cyclotron frequency for a magnetic field strength of B nT. Thus for a source at the equatorial plane in a dipole magnetic field, the radiation is beamed at angles $\alpha = (\pi/2 - \beta)$ to the equatorial plane, one beam to the north and one to the south. If the emitting region is distributed about the equatorial plane by $\sim \pm 2^\circ$ as would be the case if the jovian ESUH waves are the source, then the southern edge of the northern beam would be expected to derive from the most southerly source, that is from -2° magnetic latitude. Typical ray paths derived from ray-tracing calculations in a realistic Io torus model which clearly illustrate this configuration have been presented in Fig. 2 of ref. 10; a simplified sketch is included here in Fig. 2. Had Voyager's excursions in latitude about the equatorial plane been greater, then, according to the theory, they would have observed the high-latitude edges of the beams which, in the case of the northern hemisphere beam, would have derived from the most northerly source, that is from $+2^\circ$ magnetic latitude (see Fig. 2). This effect, where Voyagers appeared to pass right through the beams, has been observed on both Voyager 1 and 2 when outbound at 04:00 and 03:00 LT respectively⁶ despite their being at lower jovian latitudes ($\sim 5^\circ$) than when inbound, indicating that at these local times the beams make smaller angles to the magnetic equatorial plane and have a smaller width, possibly due to a smaller source dimension. These sources of bKOM observed by Voyagers when outbound are under study and will be reported separately.

The theory predicts emission of bKOM at the plasma frequency f_p at angles $\beta = \arctan(f_p/f_c)^{1/2}$ with respect to the magnetic field line. Thus assuming a source position at radial distance R_s and magnetic latitude λ_{ms} , the magnetic field magnitude and direction at that point can be computed¹¹. Furthermore,

assuming a value for f_p within the range of bKOM frequencies observed by Voyager, the theoretical direction of the beam at this frequency can also be computed. If the beam does not pass through the position at which the spacecraft lies when that particular frequency is observed, then the assumed source location is incorrect and another value of R_s is tried until one iterates onto a location that fits. If this technique is applied for all the bKOM frequencies using the corresponding positions of Voyager when they are observed, a radial profile of f_p and therefore of N_e in the Io torus can be constructed. This technique has been applied to terrestrial^{12–15} and saturnian¹⁶ myriametric radiations where it could be assumed that the magnetic fields are dipolar, and the source positions could be determined analytically. Jupiter's magnetic field near the equatorial plane at the Io torus and beyond is complicated by the presence of the equatorial current sheet¹¹ and hence the need for numerical iteration to locate the sources. Voyager 2 inbound data are used because the spacecraft's highest latitude resulted in excursions in and out of the beam during a relatively large number of consecutive jovian rotations. Results from Voyager 1 inbound and Voyagers 1 and 2 outbound will form the basis of a study whose results will be reported separately.

It is assumed, as stated above, that the southern edge of the northern bKOM beam derives from a source at $\lambda_{ms} = -2^\circ$. The positions of Voyager 2 are determined when the leading edge of bKOM at different frequencies is observed. The longitude of Voyager 2 at this time is $< 202^\circ$ CML. This is measured for the nine consecutive jovian rotations when the spacecraft was inbound from 95 R_J to 55 R_J . The Io torus plasma density profiles obtained by the remote-sensing technique described above are shown in Fig. 3. Also shown, for comparison, are the profiles determined *in situ* by two experiments^{7,17} on Voyager 1 which

Fig. 2 Schematic of bKOM beaming from the Io plasma torus. Dotted lines represent isoionic contours labelled by plasma frequency f_p . Dot-dashed line is approximate $f_p = 200$ kHz contour determined *in situ*¹ by Voyager 1. Solid line depicts where upper hybrid resonance frequency $f_{UHR} = 100$ kHz.



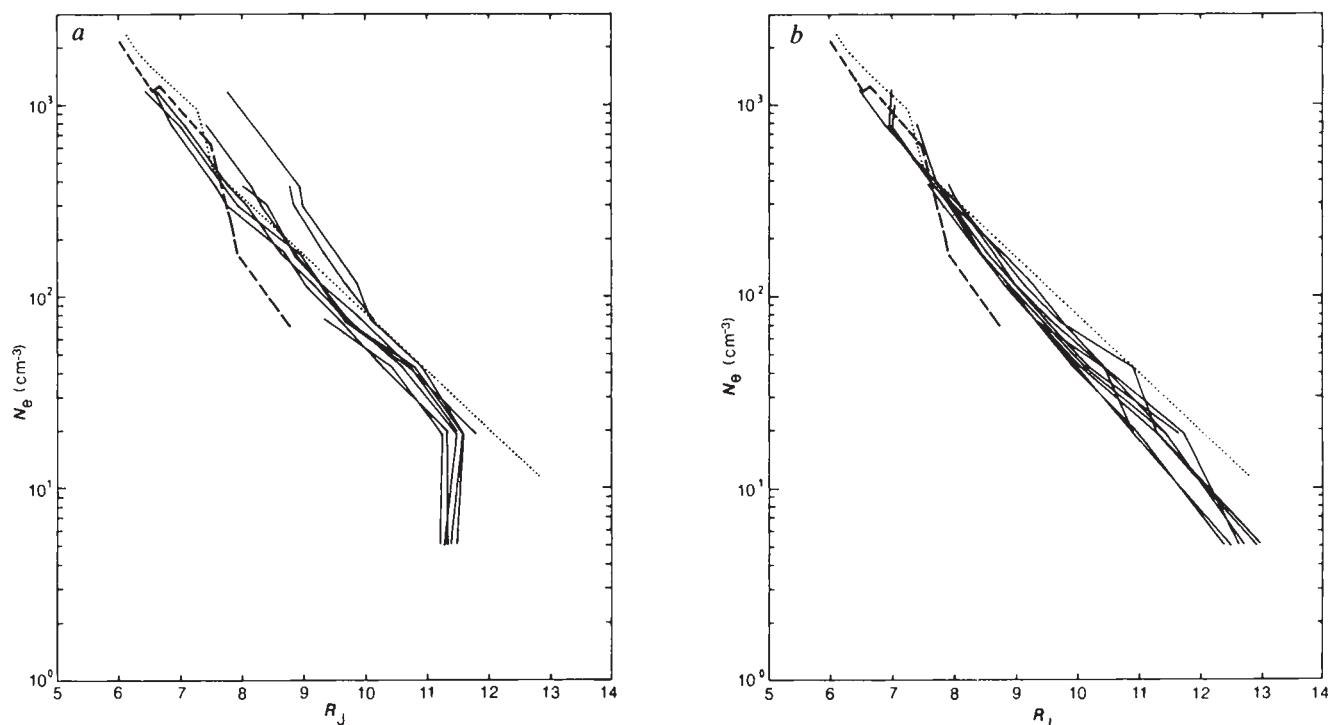


Fig. 3 *a*, Plasma density profiles of the Io torus obtained by remote sensing using the start times of bKOM observed during nine jovian rotations when Voyager 2 was inbound from 95 R_J to 55 R_J . All sources are assumed to be at magnetic latitude $\lambda_{ms} = -2^\circ$. The dashed line represents the profile determined from *in situ* Voyager 1 measurements by the plasma science PLS experiment¹⁷ and the dotted line is derived from measurements of the upper hybrid frequency by the planetary radio astronomy PRA and plasma wave science PWS experiments⁷. *b*, As for *a*, but using the end times of bKOM and assuming all sources are at magnetic latitude $\lambda_{ms} = -1^\circ$.

had traversed the torus 4 months previously. The remotely-sensed profiles are remarkably consistent within themselves considering that there were temporal variations in the torus over the 5 days involved, during which Io made nearly three rotations around Jupiter. The anomalous results at the lowest frequency are possibly due to a wave refraction effect between the source and spacecraft and this is under investigation. The agreement with the Voyager 1 *in situ* profiles is also remarkable considering the time and CML differences between the two sets of measurements. The same remote-sensing exercise has been performed using the trailing edge of the bKOM emissions, again observed in the northern hemisphere but with Voyager 2 moving down in latitude towards the equatorial plane. Figure 4 shows the profiles obtained for a source latitude of $\lambda_{ms} = -1^\circ$ which was found to give better agreement with the *in situ* profiles than $\lambda_{ms} = -2^\circ$. This difference may be due to small differences in the widths or positions of the source regions with CML since these results correspond to longitudes $>202^\circ$ CML; no *in situ* results exist for the variation with CML of the latitudinal width of the source ESUH waves, although certain asymmetries with respect to the equatorial plane have been observed¹⁸. The difference may also arise from the fact that the centrifugal equator, which is the symmetry plane for the Io torus plasma, has a magnetic latitude which differs by $\sim 1^\circ$ between the two cases due to their CML differences¹⁹.

Note that the slopes of the remotely-sensed profiles differ only slightly from that of the *in situ* profile determined from the observations of upper hybrid waves (dotted line). If one had assumed that the source latitude decreased by $\approx 0.5^\circ$ with increasing distance from 8 R_J to 13 R_J the slopes would be identical. Such a decrease in the width of the ESUH emission regions with distance has been observed¹⁸.

Because the beam widths depend to a first approximation on the latitudinal dimensions of the source, further computations show that for Voyager 2 to have observed the top edge of the northern beam when inbound on the dayside, the source ESUH

waves would have had to be restricted to within $\sim \lambda_{ms} \approx 0.5^\circ$ of the equatorial plane and even then the effect would only have been visible at the higher frequencies. Alternatively, had Voyager 2's excursion into the northern hemisphere been $> \lambda_m \approx 20\text{--}25^\circ$ it would have observed the whole beam at all frequencies.

In conclusion, bearing in mind the temporal and longitudinal variations of the torus plasma density and of the current sheet magnetic field which has been assumed constant in the model used, it is difficult to imagine that the remotely-sensed profiles obtained using data from a spacecraft $> 3 \times 10^6$ km from the source could be so consistently similar to the *in situ* profiles unless the linear theory proposed is valid.

I thank Yolande Leblanc, Observatoire de Paris, for providing the PRA data and for much advice and information, Jack Connerney, NASA Goddard Space Flight Center, for supplying the magnetic field program and Fran Bagenal, Imperial College, London, for helpful discussions.

Received 27 May; accepted 4 August 1986.

1. Warwick, J. W. *et al.* *Science* **204**, 995-998 (1979).
2. Warwick, J. W. *et al.* *Science* **206**, 991-995 (1979).
3. Kurth, W. S., Gurnett, D. A. & Scarf, F. L. *Geophys. Res. Lett.* **7**, 61-64 (1980).
4. Jones, D. *Nature* **288**, 225-229 (1980).
5. Desch, M. D. & Kaiser, M. L. *J. geophys. Res.* **85**, 4248-4256 (1980).
6. Leblanc, Y. & Daigne, G. *J. geophys. Res.* **90**, 12073-12080 (1985).
7. Barbosa, D. D. & Kurth, W. S. *J. geophys. Res.* **85**, 6729-6742 (1980).
8. Oya, H. *Planet. Space Sci.* **22**, 687-708 (1974).
9. Jones, D. *Nature* **260**, 686-689 (1976).
10. Jones, D. *Adv. Space Res.* **1**, 333-336 (1981).
11. Acuna, M. H., Behannon, K. W. & Connerney, J. E. P. *Physics of the Jovian Magnetosphere* (ed. Dessler, A. J.) 1-50 (Cambridge University Press, 1983).
12. Jones, D. *Nature* **294**, 728-730 (1981).
13. Jones, D. *Planet. Space Sci.* **30**, 399-410 (1982).
14. Jones, D. *J. Geophys. Res.* **85**, 158-166 (1983).
15. Budden, K. G. & Jones, D. *Comparative Study of Magnetospheric Systems* (ed. CNES) 563-580 (Cepaude-Editions, Toulouse, 1986).
16. Jones, D. *Nature* **306**, 453-456 (1983).
17. Bagenal, F. & Sullivan, J. D. *J. geophys. Res.* **86**, 8447-8466 (1981).
18. Kurth, W. S., Barbosa, D. D., Gurnett, D. A. & Scarf, F. L. *Geophys. Res. Lett.* **7**, 57-60 (1980).
19. Dessler, A. J. *Physics of the Jovian Magnetosphere* (ed. Dessler, A. J.) 498-504 (Cambridge University Press, 1983).

An icy-glue model of cometary nuclei

Tamas I. Gombosi* & Harry L. F. Houppis*

Central Research Institute for Physics, Hungarian Academy of Sciences, H-1525 Budapest 114, POB 49, Hungary

Since 1950 a number of models have been proposed to explain the observations of comets. For the most part, the icy conglomerate model of Whipple¹ has been the standard which others have followed. In light of the recent images made of comet Halley, by the Vega and Giotto spacecraft, we suggest here that the nucleus is composed of rather large (tens of centimetres to hundreds of metres) porous refractory boulders, 'cemented' together with an ice-dust grain mix of the icy conglomerate type. This model retains the advantages of the icy conglomerate formalism on a local level, and also introduces a new global framework. We discuss the model in terms of the improving knowledge about comets, the creation of the nucleus in the cosmogonic sense, and expectations of the future analysis of data and observations.

As well as explaining the production rate of various chemical species in the cometary coma, previous nucleus modelling has been constrained by the following observational facts: the light-curve of a comet is asymmetric with respect to perihelion; most short- and intermediate-period comets ($P < 200$ yr) 'turn on' inside ~ 3 AU; some new comets show gas activity at very large heliocentric distances (5–10 AU); and comets have been known to split. In the past few years, a number of important observations have been made which further constrain a nucleus model. Radar observations of comet IRAS-Araki-Alcock² have shown that the nucleus is highly non-spherical, that the surface is rough on a scale larger than the radar wavelength (3.54 and 12.9 cm), and that there is an expanding 'skirt' of non-gravitationally bound material with individual radii larger than a few centimetres. The clumpy nature of OH emissions in the atmosphere of comet Halley³ is suggestive of extended sublimation sources.

These most recent observations, together with an assortment of other evidence, led Weissmann⁴ to postulate a primordial rubble pile of small icy conglomerates. Although this model, and similarly that of Donn *et al.*⁵, can also explain the splitting of comets as the 'breaking free of weakly bonded fragments', the recent images of the comet Halley nucleus by the Vega and Giotto spacecraft^{6–8} suggest another way of looking at this process as well as the nucleus as a whole. The shape of the Halley nucleus is quite non-spherical ($\sim 7.5 \times 7.5 \times 15$ km³) with altitude variations of ~ 100 m. About seven major jets, some of them grouped into bundles (the source diameter of each bundle being ~ 3 km, whereas that of the single dust jets appears to be ~ 500 m) indicate an origin on the sunlit side of the nucleus and point in the general direction of the Sun. The source area of all the active jets is estimated to be $\sim 10\%$ of the total surface area. The rest of the surface appears to be inactive, although a low, undetected dust activity cannot be dismissed. Furthermore, the surface has a rather uniform albedo of a few per cent, indicative of a porous dust coating with little or no ice; this is also consistent with the infrared measurements of an average surface temperature over 300 K.

These latest results from the spacecraft encounters with comet Halley lead us to propose a new model of the nucleus. We suggest that the nucleus is composed of rather large (tens of centimetres to hundreds of metres) porous refractory boulders (similar in composition to the outer asteroid belt objects and the irregular outer satellites of the giant planets; that is, hydrous silicates containing complex carbon compounds) 'cemented' together with an ice-dust grain mix of the icy conglomerate type ('Whipple glue'). As the nucleus evolves with repeated passages

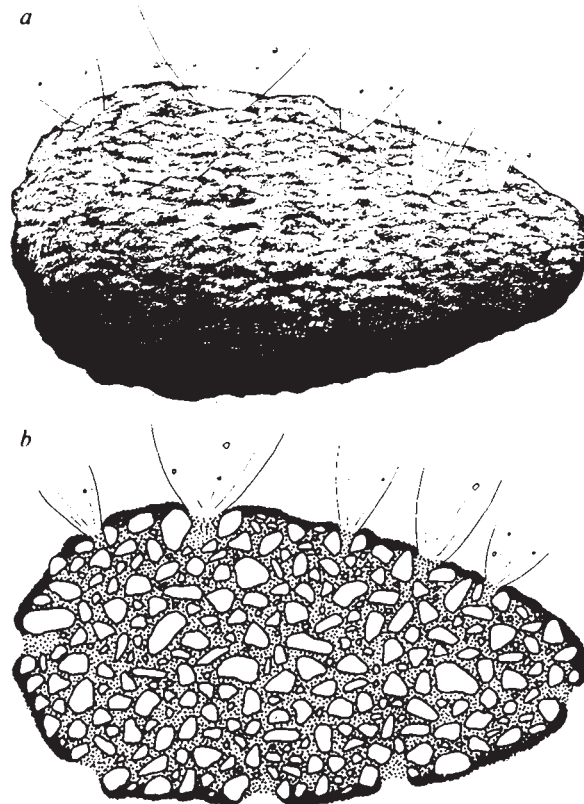


Fig. 1 *a*, Three-dimensional perspective of a comet-Halley-shaped icy-glue cometary nucleus showing active and inactive macropores, release of small boulders, surface altitude variations of porous refractory framework imbedded in Whipple glue, outer dust covering, and penetrations of the form surface by potentially active macropores.

close to the Sun, those regions above the largest boulders, which may be only a few metres thick, become depleted in their volatile ices and hence inactive. We call these regions dust plates, and they may be only a few centimetres thick. On the other hand, the regions between the largest boulders, where there are the smaller boulders and the ice-dust mix, may penetrate quite deep into the nucleus (50–100 m) and hence are not likely to become inactive for several (perhaps 50–100) orbits around the Sun. We call these regions macropores to distinguish them from the micropores in the Whipple glue. It is from here that the jets originate. Figure 1 shows conceptual drawing of such a nucleus.

This model explains in a natural way the above observations. The general shape of the nucleus and its surface altitude variations are a direct consequence of the low self-gravity and the large boulder framework, whereas the uniform porous dust covering is a result of the volatile ices receding below the surface and the creation of dust plates.

Jets, as the major sources of cometary activity, are the result of a group of macropores coming into view of the Sun. Depending on the placement of the boulders, the jets may appear as line sources delineating the shape of the large boulders near the surface or as point sources for a more evolved surface topology. These appearances are reminiscent of the surface regions sketched by Sekanina and Larson^{9,10}. Furthermore, the latest analysis of the Vega images indicates the existence of both linear and point sources on comet Halley (B. A. Smith, personal communication).

The percentage of comet Halley's surface capable of significant dust emission can be estimated from the observed jet activity. Using a total active area of ~ 50 km² as seen by Keller *et al.*⁶ for the Halley nucleus, one can conclude that as much as 44% of the entire surface is capable of jet activity.

Typically, we expect the width of the smallest dimension of

* Present address: Space Physics Research Laboratory, Space Research Building, University of Michigan, Ann Arbor, Michigan 48109, USA (T.I.G.); Centre of Astrophysics and Space Science, University of California, San Diego, California 92093, USA (H.L.F.H.).

a jet at the surface to be no more than a few hundred metres. However, a few hundred metres from the surface the jet will have expanded laterally due to the existence of a gas-pressure gradient between the site above the macropore and that of the surrounding dust plates¹¹.

The literature often suggests that jets may originate from fissures due to thermal stress, which have opened on the surface. The freshly exposed ice then provides the source for the jets. Our contention is that although the surface may crack, the resulting excess production will be of an outburst nature rather than a steady jet. The reason is that exposed ice will quickly recede to form a dust mantle (only a few tenths of a centimetre are needed) and return the excessively active surface to that of a diffusion process. Thus, although fissures and unglued small boulders may account for outbursts, the macropores are responsible for the steady jet activity.

Splitting of comets is now explained by pieces of the framework becoming unglued; that is, the Whipple glue around a particular boulder has been sublimed and entrained to such a point, that the boulder eventually breaks free. This process which is obviously random with respect to the comet's orbital parameters and heliocentric distance, explains why more splitting fragments are a small fraction of the original nucleus, separating with a small relative velocity and disappearing visually in ~ 100 days, and allows for 'flaring' activity as the new uncovered Whipple glue is exposed to the solar environment. The decrease in splitting frequency as a comet spends more time in the inner Solar System may be due to such processes as densification of the ice-dust grain mix¹², and the recondensation near the surface of volatile ices which have sublimed from the comet interior¹³. We also suggest that outbursts from comets may be in part the result of the much smaller boulders becoming unglued. We note that a major difference between our mechanism and that of the rubble-pile model is that splitting of the icy glue nucleus is a natural consequence of the structural make-up, whereas the rubble-pile theory must invoke 'weakly bonded fragments' and the necessity for a heat wave to have some *ad hoc* preference to propagate along the interfaces between them.

The activity of some 'new' comets at large heliocentric distance may also be partially explained by postulating that these comets are indeed new to the inner Solar System, with the entire surface capable of dust and ice production, and that comets are probably significantly larger than previously thought. Thus, it is not until the inactive dust plates begin to predominate after a few orbits that activity at large heliocentric distances begins to subside. This idea may of course be used in conjunction with the chemical differentiation model¹⁴ to account for the 'tapering off' of activity at smaller heliocentric distances. The asymmetric behaviour of cometary light curves about perihelion can be explained in part by the location of the active macropores.

For the creation of an icy-glue nucleus we suggest the following model. The large porous boulders first agglomerated in the pre-Jupiter region. The growing Jupiter and Saturn protoplanets caused the diffusion of a substantial amount of this material through the pre-Uranus-Neptune zone. Because agglomeration of the original influx of material into this outer zone was quite slow, as suggested by a number of theories¹⁵, the boulders arrived to find the ice and dust still in the small-grain phase and somewhat extended about the ecliptic plane. The boulders and grains then came together to form the icy-glue nuclei described previously, with the Uranus and Neptune protoplanets ejecting them to the Oort cloud.

Although this model is only illustrative and needs to be subjected to a proper dynamical analysis, we point out an important difference between the icy-glue complex and the icy conglomerate models. On the one hand, the icy conglomerate model has its basic building blocks constructed from both dust and ice grains, although the mixture may be heterogeneous relative to the centre of the nucleus or the individual building blocks. On the other hand the icy-glue model is built from two

different types of basic units: porous refractory boulders with little or no initial volatile content, and grains of dust and ice. This fundamental difference may prove useful in constructing a cosmogonic theory of the Solar System.

We thank Drs Bradford Smith, Akos Kőrösmezey, Steven M. Larson and Károly Szegő for useful discussions, Carole Turner for the careful editing of this manuscript and Diane McGregor for the artwork of Fig. 1. Part of this work was supported by NSF grant AST-8605994.

Received 15 May; accepted 11 August 1986.

1. Whipple, F. L. *Astrophys J.* **111**, 375-349 (1950).
2. Goldstein, R. M., Jurgens, R. F. & Sekanina, Z. *Astr. J.* **89**, 1745-1754 (1984).
3. dePater, I., P. & Schenewerk, M. *Pap. presented at Conf. Am. Geophys. Un.*, San Francisco, December 1985; *IAU Circ.* No. 4142 (1985).
4. Weissman, P. *Nature* **320**, 242-244 (1986).
5. Donn, B., Hughes, D. W. & Daniels, P. A. *Bull. Am. astr. Soc.* **17**, 520 (1985).
6. Keller, U. *Pap. presented at Max Planck Institut für Aeronomie Seminar*, Katlenburg-Lindau, W. Germany, 2 April 1986; Keller U. *et al Nature* **321**, 320-326 (1986).
7. Sagdeev, R. Z. *et al Nature* **321**, 259-262 (1986).
8. Combes, M. *et al Nature* **321**, 266-268 (1986).
9. Sekanina, Z. *Astr. J.* **86**, 1741-1773 (1981).
10. Sekanina, Z. & Larson S. *Nature* **321**, 357-361 (1986) *Astr. J.* (submitted).
11. Kitamura, Y. *Icarus* (submitted).
12. Smoluchowski, R., Marie M. & McWilliam, A. *Earth, Moon & Planets* **30**, 281-288 (1984).
13. Whipple, F. L. & Stefanik, R. P. in *Nature et Origine des Comètes*, 13th Liège Symp., 33-51 (1966).
14. Houpis, H. L. F. Ip, W.-H., & Mendis, D. A. *Astrophys. J.* **245**, 654-667 (1985).
15. *Protostars and Planets II* (eds Black, D. C. & Matthews, M. S.) Part VII, pp. 959-1155 (University of Arizona Press, 1985).

Distribution of energetic particles near interplanetary shocks

A. Hewish* & S. Bravo†

* Mullard Radio Astronomy Observatory, Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

† Instituto de Geofísica, Ciudad Universitaria, Código 04510, Mexico D.F.

The acceleration of charged particles by collisionless shocks is a process of considerable importance in astrophysics, but the physical mechanism is not yet understood in detail. Energetic particles near interplanetary shocks may be studied *in situ* by spacecraft, but little is known about their distribution across the shock surface because of inadequate knowledge of the location of the sampling point in relation to the whole disturbance. A ground-based method has been developed² which enables shock-associated disturbances to be mapped in three dimensions, and we have used it, in combination with near-Earth spacecraft measurements, to study how the particle flux varies for shocks travelling at different angles to the Sun-Earth line. We have observed a pronounced east-west asymmetry which indicates that the inclination of the interplanetary magnetic field to the shock surface must be a potent factor in the acceleration process.

The method of mapping the disturbances was based on daily observations of a large grid of radio sources, the scintillation of which was related to the mean plasma density along many different lines of sight¹. Examples of the structures of large-scale disturbances revealed by this method have been described previously²⁻⁴. It was found that strong shocks typically occurred at the front of enhanced density shells which subtended a solid angle of $\sim \pi/2$ sr at the Sun. The shells were compression zones, caused by sudden eruptions of fast solar wind which lasted for several days, and they could always be traced back to sources coincident with coronal holes⁵. Among the 96 disturbances detected during August 1978-September 1979, 73 were classified as 'erupting streams'; the remainder were co-rotating interaction regions of longer duration⁶. All significant shocks (shock strength ≥ 1.5) recorded by near-Earth spacecraft were accompanied by compression zones that were mapped by our method.

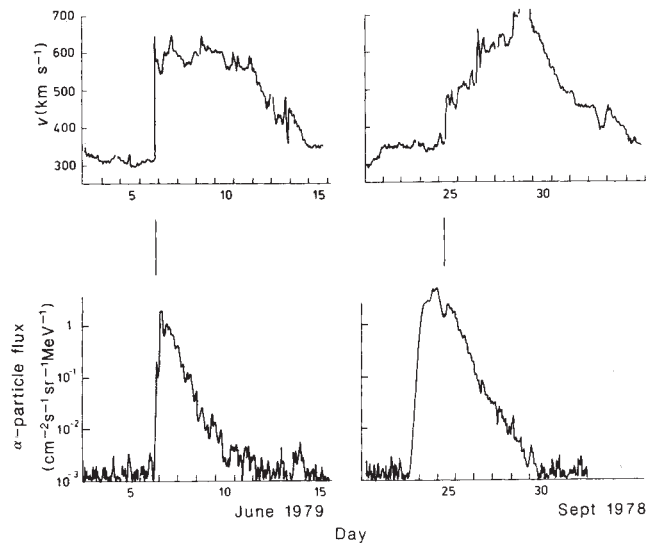


Fig. 1 Temporal variations of solar wind speed (v) and α -particle flux in the range 9–70 MeV associated with two strong interplanetary shocks. Vertical line, time of shock passage at the Earth.

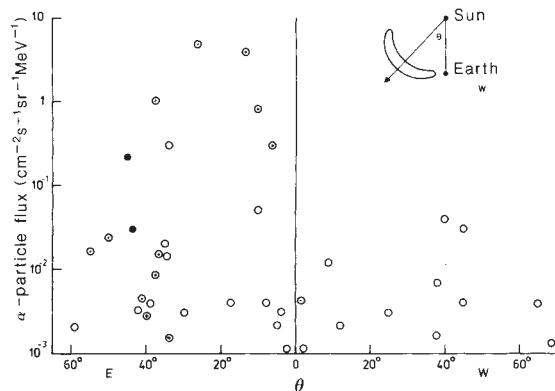


Fig. 2 Peak α -particle flux observed for shocks travelling at different angles, θ , to the Sun-Earth line. Shock strengths: ●, ≥ 3 ; ◐, ≥ 2 ; ○, < 2 .

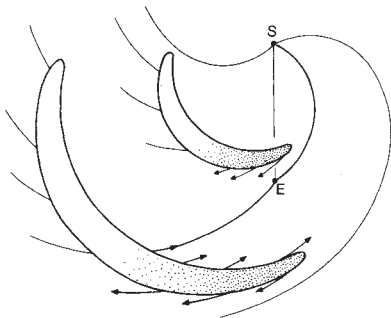


Fig. 3 Schematic diagram showing the regions of greatest particle acceleration when the shock disturbance has reached distances from the Sun (S) 1 and 2 AU. The spiral magnetic field lines correspond to a solar wind speed of 350 km s^{-1} . Arrows denote the directions of particles moving away from the shock. The field line through the Earth (E) beyond 1 AU is calculated for a speed of 600 km s^{-1} , appropriate to the high-speed flow behind the shock.

However, we sometimes detected eruptions unaccompanied by shocks, either because the shocks were too weak or because they did not strike the Earth.

All the disturbances that were classed as erupting streams, and that also caused sudden geomagnetic storms, are listed in Table 1 together with spacecraft measurements of the shock strength. Also given are the estimated directions of travel (θ), relative to the Sun-Earth line, obtained by our method. As described previously, the angles were obtained by comparison with theoretical maps computed for simple models⁶. The accuracy of ($\sim \pm 20^\circ$) is not high, but there was usually no difficulty in placing disturbances to the east or west of the Sun-Earth line.

For the energetic particle data we used the α -particle flux in the range 9–70 MeV from the GMS/SEM satellites⁷. During the largest particle events the flux increased by a factor of $\sim 10^3$ (Fig. 1). Table 1 gives the peak flux associated with each shock, smoothed over a few hours. When the particle fluxes are plotted against the directions of travel (Fig. 2), a pronounced asymmetry is evident. Only shock disturbances centred on directions to the east of the Sun-Earth line produced the largest enhancements of flux. The background flux was typically $\sim 2 \times 10^{-3} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \text{ MeV}^{-1}$, and enhancements greater than four times this value occurred on average only once or twice per month. At this level there were 16 events associated with shocks passing on the east, compared with 3 in the west. Above $5 \times 10^{-2} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \text{ MeV}^{-1}$, there were 8 shocks in the east and none in the west.

Table 1 Shock disturbances striking the Earth

	Date	Arrival time UT	θ	Particle flux	Shock strength
1978	18 Aug	1242	38 W	1.5×10^{-3}	1.4
	27 Aug	0246	43 E	4×10^{-3}	2.5
	5 Sep	1859	5 E	2×10^{-3}	1.6
	9 Sep	0354	38 W	7×10^{-3}	—
	11 Sep	0900	69 W	1.5×10^{-3}	—
	20 Sep	1417	30 E	3×10^{-3}	—
	25 Sep	0718	26 E	5.0	> 2
	29 Sep	0301	9 W	1.2×10^{-2}	1.8
	4 Oct	0047	37 E	1.3×10^{-2}	2.2
	9 Oct	0332	35 E	2×10^{-2}	1.5
	17 Oct	0430	4 E	3×10^{-3}	1.4
	29 Oct	1116	42 E	3.5×10^{-3}	1.7
	8 Nov	0152	59 E	2×10^{-3}	1.5
	12 Nov	0100	6 E	3×10^{-1}	2.8
	18 Nov	2320	2 E	1×10^{-3}	1.2
	18 Dec	0036	12 W	2×10^{-3}	—
	25 Dec	1212	40 E	2.8×10^{-3}	2.0
1979	4 Jan	1725	45 W	4×10^{-3}	—
	6 Jan	0032	40 W	4×10^{-2}	1.7
	9 Jan	0340	55 E	1.5×10^{-2}	2.2
	25 Jan	0139	39 E	3.8×10^{-3}	1.9
	21 Feb	0302	50 E	1.3×10^{-2}	2.2
	4 Mar	0445	35 E	1.5×10^{-2}	—
	6 Mar	0818	45 W	3×10^{-2}	1.5
	9 Mar	0808	25 W	3×10^{-3}	1.7
	17 Mar	0231	65 W	4×10^{-3}	—
	22 Mar	0826	34 E	1.3×10^{-3}	2.5
	5 Apr	0156	37 E	1.0	2.0
	24 Apr	2358	43 E	3×10^{-2}	3.3
	29 May	1851	38 E	8×10^{-3}	2.2
	6 Jun	1927	13 E	4.0	2.9
	6 Jul	1930	45 E	2×10^{-1}	3.2
	12 Jul	1240	7 E	3×10^{-3}	—
	11 Aug	1812	18 E	4×10^{-3}	1.6
	13 Aug	0639	2 W	1×10^{-3}	—
	20 Aug	0625	10 E	8×10^{-1}	2.4
	24 Aug	1457	10 E	5×10^{-2}	1.7
	29 Aug	0459	34 E	3×10^{-1}	—

The shock strengths (downstream/upstream density) were taken from refs 13, 14. The α -particle flux is in units of $\text{cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \text{ MeV}^{-1}$.

Assuming that all the shocks are similar, it follows that the particle flux at the shock surface follows the distribution shown in Fig. 2, but in the opposite sense. Thus, a shock travelling directly towards the Earth would generate a flux which peaked at 20–30° west of the Sun–Earth line.

The asymmetry can be understood if the acceleration depends on the angle between the magnetic field and the shock surface. Figure 3 shows the geometry for an idealized interplanetary field, assuming an ambient solar wind of 350 km s⁻¹. For co-rotating shocks beyond 1 AU caused by a stable high-speed stream, the maximum particle flux occurs when the shock surface is nearly parallel to the field⁸ (the 'quasi-perpendicular' case). For a spherical shock front appropriate to an erupting stream, it is clear from Fig. 3 that this condition can occur on one side of the disturbance only, and the asymmetry is in the correct sense to explain our observations. Another factor favouring the parallel field condition is the co-rotation of the stream following its sudden eruption. This will tend to increase the curvature of the shock front on the same side of the disturbance.

In addition to the asymmetry, the characteristic sharp rise and slow decay of the flux (see Fig. 1) is readily explained. Recent evidence^{9,10} supports a model in which particles are accelerated continuously near the shock and then diffuse away in both directions along the field lines. From the geometry in Fig. 3 it may be seen that for a shock passing to the east, the field line through the Earth makes connection only shortly before the shock arrives, but thereafter remains connected for several days as the disturbance propagates beyond 1 AU. The systematic eastward movement of the point of contact as the shock recedes from the Earth effectively scans the particle distribution in a different way.

In previous studies of energetic particle events, associations with solar flares have been assumed and conflicting results have been obtained. It has been deduced, for example, that flares are most effective when they occur on the west of the disk, which supports the theory that energetic particles are released at the flare site and channeled along 'well-connected' field lines to the Earth¹¹. By contrast, more recent work has suggested that the flux is greatest for shocks associated with flares on the east of the disk¹². These discrepancies are not surprising in the light of our observations that shocks are caused by eruptions generated by coronal holes and not by flares⁶. Our present conclusions are qualitatively similar to those of ref. 12, but differ in showing a larger and more sharply defined east–west asymmetry for the particle flux.

Received 23 June; accepted 24 July 1986.

1. Tappin, S. J. *Planet. Space Sci.* **34**, 93–97 (1986).
2. Gapper, G. R., Hewish, A., Purvis, A. & Duffett-Smith, P. J. *Nature* **296**, 633–636 (1982).
3. Tappin, S. J., Hewish, A. & Gapper, G. R. *Planet. Space Sci.* **31**, 1171–1176 (1983).
4. Tappin, S. J., Hewish, A. & Gapper, G. R. *Planet. Space Sci.* **32**, 1273–1281 (1984).
5. Hewish, A., Tappin, S. J. & Gapper, G. R. *Nature* **314**, 137–140 (1985).
6. Hewish, A. & Bravo, S. *Sol. Phys.* (in the press).
7. Nagai, T. & Kohno, T. *Sol. terr. envir. Res. Japan* **8**, 20–33 (1984).
8. Tsurutani, B. T., Smith, E. J., Pyle, K. R. & Simpson, J. A. *J. geophys. Res.* **87**, 7389–7404 (1982).
9. Evenson, P., Meyer, P. & Yanagita, S. *J. geophys. Res.* **87**, 625–631 (1982).
10. Sanahuja, B., Domingo, V., Wenzel, K.-P., Joselyn, J. A. & Keppler, E. *Sol. Phys.* **84**, 321–337 (1983).
11. Van Hollebeke, M. A. I., Ma Sung, L. S. & McDonald, F. B. *Sol. Phys.* **41**, 189–193 (1975).
12. Sarris, E. T., Decker, R. B. & Krimigis, S. M. *J. geophys. Res.* **90**, 3961–3965 (1985).
13. van Nes, P., Reinhard, R., Sanderson, T. R. & Wenzel, K.-P. *J. geophys. Res.* **89**, 2122–2132 (1984).
14. Cane, H. V. *J. geophys. Res.* **90**, 191–197 (1985).

Searching potential energy surfaces by simulated annealing

L. T. Wille

SERC Daresbury Laboratory, Daresbury, Warrington WA4 4AD, UK

Many problems in physics¹, chemistry, biology² and mathematics³ involve the determination of the absolute minimum of a certain multidimensional function. In most cases of practical interest this is a complicated matter, owing to the presence of local minima, and even more so because the number of local minima often increases exponentially with the problem size. Standard techniques apply local optimizers to many random initial configurations, but soon become intractable as the dimensionality increases. Here I show how the simulated annealing method can be used to guide a search towards the absolute minimum. The method is illustrated on a problem recently discussed in this journal, namely the minimum-energy configuration of equal charges confined to a sphere. This problem, although easy to visualize, can be used to simulate much of the complexity of the above problems by considering a large number of particles. In this limit several minima have been found that have previously been missed by other authors using classical techniques.

To determine the ground-state configuration of a system of interacting particles one needs to find the absolute minimum on the potential-energy surface (strictly speaking it is the free energy that has to be minimized, but one usually assumes the system to be at zero temperature). A local minimum can be found by one of many possible search techniques⁴, most of which are based on a gradient method. Starting from a certain point on the potential energy surface one takes a step in a direction determined by the line of steepest descent and repeats this process until a minimum is reached, to within a certain accuracy. The collection of all points from which the same

(local) minimum is obtained by this algorithm is called the catchment region of the minimum. This whole procedure is repeated a number of times starting from random initial configurations and produces a list of local minima, the smallest of which is taken to be the global minimum. Obviously there is no guarantee that the global minimum will be found in this way, especially if the catchment region is narrow. Alternatively, instead of taking random initial configurations one tries to deduce growth schemes from low-dimensional systems. The algorithm is then applied to configurations obtained from a set of growth rules. This procedure often produces highly symmetric configurations; the previous remarks are then even more pertinent.

Obviously a method that avoids getting trapped in local minima would increase the likelihood of detecting the global minimum. It has been shown^{5,6} that a continuous version of the simulated annealing method proposed by Kirkpatrick *et al.*⁷ for discrete problems can be used for this purpose. An overview of simulated annealing and its applications has been presented elsewhere⁸ and only a brief description will be given here. Consider the potential energy to be minimized as $E(\mathbf{r}_1, \dots, \mathbf{r}_N)$, where the $\mathbf{r}_i$, $i = 1, \dots, N$ are the positions of the particles. The simulated annealing method uses the Metropolis prescription⁹ to decide whether or not to accept a random step $\Delta \mathbf{r}_i$, $i = 1, \dots, N$. If the associated energy change $\Delta E = E(\mathbf{r}_1, \dots, \mathbf{r}_i + \Delta \mathbf{r}_i, \dots, \mathbf{r}_N) - E(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N)$ is negative, the step is accepted, otherwise it is accepted with a probability $P = \exp(-\Delta E/T)$, where T is the temperature (in reduced units). Obviously this corresponds to a random walk on the potential energy surface, where the walker feels attracted to minima. In the process of annealing the system is started in a random initial configuration at a sufficiently high temperature. Several series of steps are attempted and the temperature is then reduced by a certain factor (typically 0.75–0.9). This process is repeated until no further improvements have been made over a number of iterations. This scheme is similar to the way a crystal is grown from the melt. If the fluid is cooled too quickly (or quenched) a

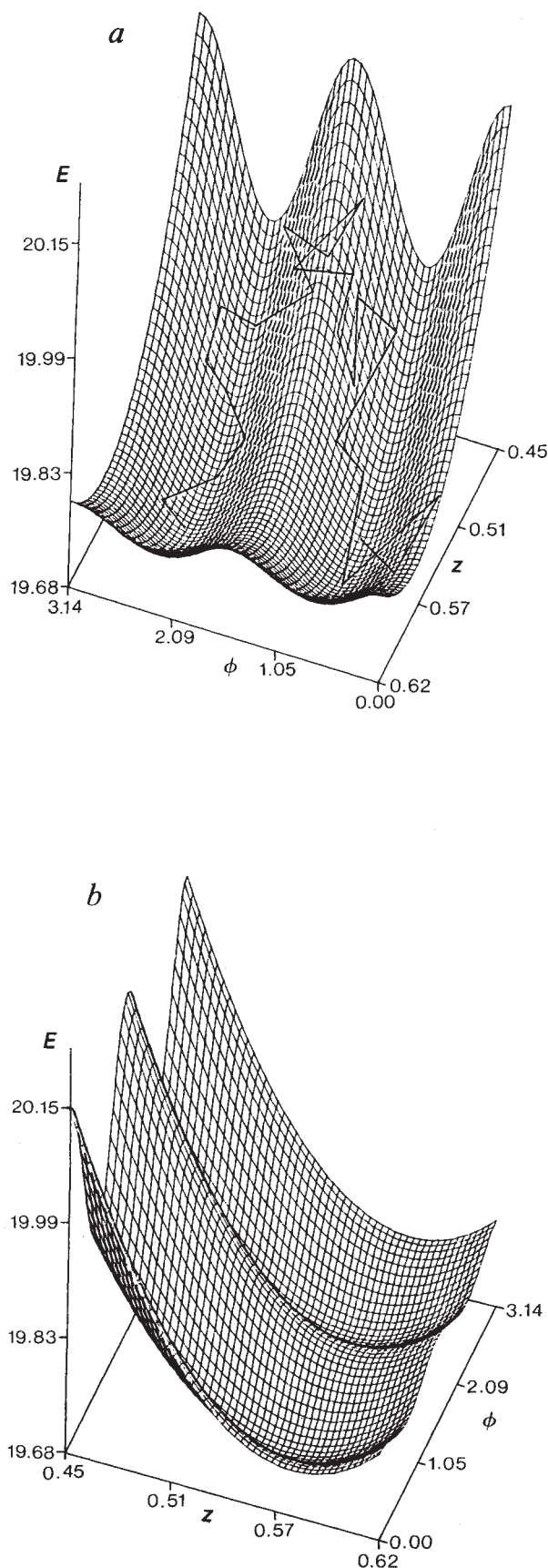


Fig. 1 The potential energy surface for an antiprismatic arrangement of charges: z is the distance of the planes from the equator and ϕ is the relative rotation angle. *a*, The random walk towards the minimum. *b*, A cubical arrangement $\phi = \frac{1}{2}\pi$, $z = 0.577$ corresponding to a saddle point on the potential energy surface.

disordered structure with many defects results. This would correspond to a local minimum. On the other hand, the true crystalline ground state can be obtained by a process of slow annealing. It was this powerful analogy which motivated Kirkpatrick *et al.* to develop the simulated annealing method⁷.

In the calculations presented here the initial temperature was taken from a short preliminary simulation and the step length was adjusted so that 25–50% of the steps were accepted. Typically 100 moves for each particle were attempted before the temperature was reduced by a factor of 0.9. However, it was found that in many cases a much faster cooling produced the same results. As soon as the energy dropped below a certain value (also taken from a preliminary simulation) a steepest-descent optimizer was called. Finally the whole process was repeated ten times for each number of particles. This precaution was taken because it was found previously⁶ that, in spite of slow cooling, trapping in a low-lying local minimum can still occur. The use of a local optimizer in the later stages of the simulation is also an innovation and gives rise to a great reduction in computation time.

Recently there has been renewed interest in the minimum-energy configurations of systems of charged particles confined to a circle or a sphere^{10–17}. The problem of the distribution of equal point charges on a sphere is long standing; it was investigated at the beginning of this century by J. J. Thomson and D. Hilbert and later by many mathematicians (for a historical review see ref. 18). Very little analytical progress has been made and most of the later work^{19–21} uses numerical methods to find the global minimum for this and other extremal problems over a spherical surface. It has been shown²² that if the configuration possesses rotational symmetry about the diameter through every particle, this configuration is in equilibrium whatever the force law is, provided that it depends only on the distance between the particles. In this work the charges are assumed to interact via a Coulomb potential $V = 1/r$. It is then known that there are no minimum-energy configurations with charges inside the sphere because the potential is harmonic, that is, it satisfies the Laplace equation^{23,24}. Hence the particles can be restricted to move on the surface of the sphere during the simulation, a simplification that would no longer be possible for a potential of the form $1/r^n$ ($n > 1$).

Table 1 shows the minimum-energy configurations for N unit charges ($5 \leq N \leq 20$) confined to a sphere with unit radius, determined by the simulated annealing method. These configurations have been rotated so that the major symmetry axis lies along the north-south direction. The charges are grouped in planes parallel to the equator. In the planes the charges form polygons which are rotated relative to one another. A notation such as $15^3 1$ (height 0.610) means that there is one charge at the north pole, 5 charges at a height 0.610 (in units of the sphere radius), 5 in the equator plane, 5 at a height -0.610 and one at the south pole. The surprising fact that the charges tend to form parallel polygons has been noticed by previous authors. However, this is not a universal feature as is illustrated by the low-symmetry solution for $N = 11$, missed in ref. 17, but found by other authors^{19,21}. Other improvements over the results in ref. 17 for $N = 13$ and 15 have also been given before²¹. However, the present results disagree with published results for $N = 14$, which according to ref. 21 should form a $14^3 1$ configuration (with an energy of 69.496), but which is found here and in ref. 17 to form a $16^2 1$ configuration. Finally, in the range $N \leq 17$ –20 which was investigated for the first time in ref. 17, improvements have been found for $N = 19$ and $N = 20$.

Figure 1 illustrates visually how simulated annealing works, by showing the results of an actual simulation for $N = 8$. It is known that when $N = 8$ the charges do not occupy the corners of a cube, and in general the Platonic solids do not necessarily produce the minimum-energy configuration. Although Leech's analysis¹² shows that charges placed on the corners of a regular polyhedron are in equilibrium, this is not in general a stable

Table 1 Minimum energy configuration for N charges confined to a sphere, interacting via a Coulomb potential

N	Configuration	Heights	Energy
5	1 3 1		6.4747
6	1 4 1		9.9853
7	1 5 1		14.453
8	4 ²	0.560	19.675
9	3 ³	0.704	25.760
10	1 4 ² 1	0.423	32.717
11	1 2 4 2 ²	0.515, 0.168, -0.553, -0.806	40.596
12	1 5 ² 1	0.447	49.165
13	1 2 ² 4 2 ²	0.611, 0.501, -0.104, -0.545 -0.863	58.853
14	1 6 ² 1	0.455	69.306
15	3 ⁵	0.827, 0.407	80.670
16	4 ⁴	0.785, 0.197	92.920
17	1 5 ³ 1	0.610	106.05
18	1 4 ⁴ 1	0.675, 0.203	120.08
19	1 4 2 4 2 ² 4	0.638, 0.540, -0.071, -0.197 -0.333, -0.823	135.09
20	1 3 6 ² 3 1	0.693, 0.542	150.88

Results obtained by the simulated annealing method that are believed to correspond to the global minima on the potential energy surface.

equilibrium. A cubical arrangement, for example, corresponds to a saddle-point on the potential energy surface. Consequently it is possible to deform the cube continuously to the true ground state²⁰, a case of spontaneous symmetry breaking. To show how the method works, I use the following simulation. Consider the restricted problem where two squares of charges lie in parallel planes a distance z above and below the equator and with relative rotation angle ϕ . The cube corresponds to $z=1/\sqrt{3}$ ($=0.577$) and $\phi=0$, and the true ground state is $z=0.560$, $\phi=\pi/4$. Figure 1 shows the potential energy surface $E(z, \phi)$ for this constrained geometry, and the random walk towards the minimum (Fig. 1a). The potential energy $E(z, \phi)$ is a periodic function of ϕ , with period $\pi/2$; hence there are four equivalent minima, two of which are shown in Fig. 1. The simulation started in the right-hand side in Fig. 1a, at $T=0.1$, attempted one move ($\Delta z, \Delta \phi$) (Δz and $\Delta \phi$ are distributed uniformly in $[-0.1, 0.1]$ and $[-0.5, 0.5]$ respectively) after which the temperature was reduced by a factor of 0.9. Initially (at 'high' temperatures) almost all steps are accepted, then there is an intermediate regime where the walker is wandering on a 'ridge', and finally it walks down the 'valley' towards the minimum. Note that in a Monte Carlo simulated annealing the walker tunnels through barriers, rather than surmounting them as in a molecular dynamics simulation.

Simulated annealing is a powerful way of finding low-lying minima on a potential-energy surface, in particular for cases where steepest descent methods become too time-consuming. This has been illustrated by presenting several improvements over previous work for the three-dimensional charge-confinement problem, which is of mathematical interest¹⁸ and may be relevant to several fields of physics¹⁷. Moreover, the method is very general and can easily be applied to similar problems. Results for clusters of atoms interacting under two-body Lennard-Jones forces¹ will be presented elsewhere. Even in cases where the interactions between particles cannot be expressed as a two-body force, as in quantum chemistry or solid state physics, a generalization of the method, following Car and Parrinello²⁵, may be possible.

Received 6 June; accepted 18 August 1986.

- Hoare, M. R. & McInnes, J. *Adv. Phys.* **32**, 791-821 (1983).
- Furois-Corbin, S. & Pullman, A. *Chem. Phys. Lett.* **123**, 305-310 (1986).
- Clare, B. W. & Kepert, B. L. *Proc. R. Soc. A* **405**, 329-344 (1986).
- Powell, M. J. D. in *Nonlinear Optimization 1981* (Academic, New York, 1982).
- Vanderbilt, D. & Louie, S. G. *J. comp. Phys.* **56**, 259-271 (1984).
- Wille, L. T. & Vennik, J. *J. Phys. A* **18**, L1113-L1117 (1985); *J. Phys. A* **19**, 1983 (1986).

- Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. *Science* **220**, 671-680 (1983).
- Wille, L. T. *CCP5 Inform. Q.* **20**, 19-31 (Daresbury Laboratory).
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. & Teller, E. *J. chem. Phys.* **21**, 1087-1092 (1953).
- Berezin, A. A. *Nature* **315**, 104 (1985).
- MacGowan, D. *Nature* **315**, 635 (1985).
- Nityanada, R. *Nature* **316**, 301 (1985).
- Cormack, A. M. *Nature* **316**, 301-302 (1985).
- Queen, N. M. *Nature* **317**, 208 (1985).
- Aspden, H. *Nature* **319**, 8 (1986).
- Callen, M. G., Kiang, D. & Tindall, D. A. *Nature* **319**, 454 (1986).
- Munera, H. A. *Nature* **320**, 597-600 (1986).
- Whyte, L. L. *Am. math. Monthly* **59**, 606-611 (1952).
- Cohn, H. *Math. Tab. natn. Res. Coun. Wash.* **10**, 117-120 (1956).
- Goldberg, M. *Maths Comput.* **23**, 785-876 (1969).
- Melnik, T. W., Knop, O. & Smith, W. R. *Can. J. Chem.* **55**, 1745-1761 (1977).
- Leech, J. *Math. Gaz.* **41**, 81-90 (1957).
- Kellog, O. D. *Foundations of Potential Theory* (Springer, Berlin, 1967).
- Helms, L. L. *Introduction to Potential Theory* (Wiley, New York, 1969).
- Car, R. & Parrinello, M. *Phys. Rev. Lett.* **55**, 2471-2474 (1985).

Large AlCuLi single quasicrystals with triacontahedral solidification morphology

B. Dubost, J.-M. Lang, M. Tanaka, P. Sainfort & M. Audier*

Cegedur-Pechiney, Centre de Recherche et Développement, B.P.27, 38340 Voreppe, France

* Permanent address: Laboratoire de Thermodynamique et Physico-Chimie Métallurgiques, UA CNRS 29, Domaine-Universitaire, B.P. 75, 38402 Saint Martin d'Hères, France

The discovery of icosahedral symmetry in rapidly solidified Al-Mn alloys¹ has stimulated considerable research activity mainly because it appeared to violate the classical rules of crystallography. Hence, it left the field open for numerous reports on rapidly-solidified phases with icosahedral symmetry in aluminium based alloys. However, a challenging goal for materials scientists has always been the fabrication of large quasicrystals that would allow further investigation of the crystallographic and physical properties of this new form of material. This paper reports the first evidence of triacontahedral solidification morphology of large quasicrystalline dendrites within an as-cast Al-Cu-Li alloy with composition close to that of the Al₆CuLi₃ phase.

The first evidence of stable intermetallic phases with icosahedral symmetry, in the aluminium-lithium-copper-magnesium system², can be considered as an important step towards the fabrication of large quasicrystals. The intermetallic Al₆CuLi₃ compound, designated T2 (ref. 3) has been successfully cast into large dendrites surrounded by a residual eutectic, and nucleated and grown as a fine precipitate in an aluminium-rich matrix⁴.

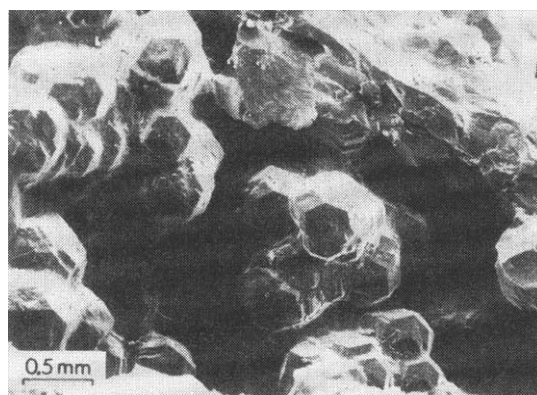
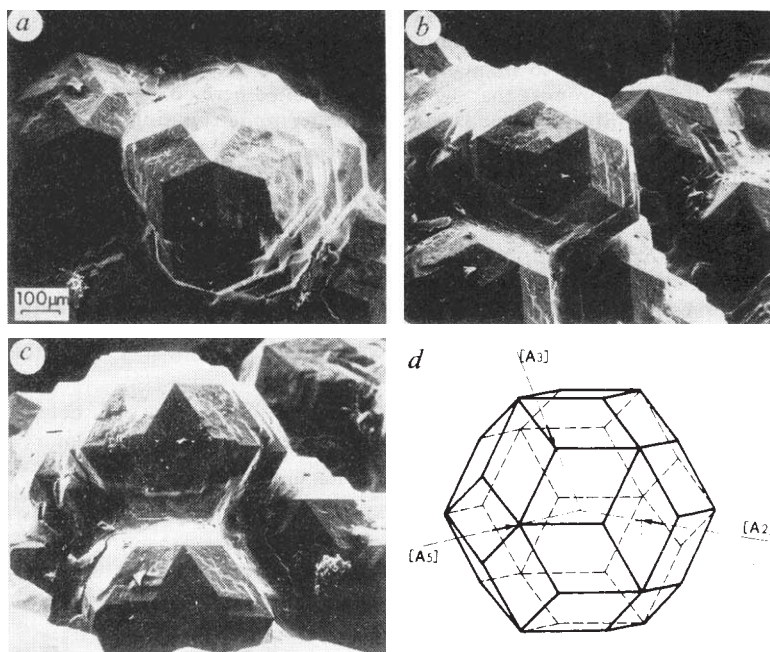


Fig. 1 Scanning electron micrograph in secondary electron mode showing the macroscopic morphology of the T2-Al₆CuLi₃ single quasicrystalline dendrites. Each dendrite consists of stacked triacontahedra with the same orientation.

Fig. 2 Triacontahedral solidification morphology of a $T2-Al_6CuLi_3$ single quasicrystal viewed along a 5-fold axis (*a*), a 3-fold axis (*b*) and a 2-fold axis (*c*). These types of symmetry axes are noted on the drawing of a triacontahedron in (*d*).



These T2 phases revealed an icosahedral structure in orientation relationships with the face centred cubic matrix⁴.

However, the direct production of single-phase intermetallic alloys with T2 composition proved difficult because of the occurrence of the cubic $R-Al_5CuLi_3$ intermetallic compound having a composition very close to that of $T2-Al_6CuLi_3$ (ref. 3). An inadequate alloy composition would then lead to a duplex matrix structure consisting of an R phase polyhedral core surrounded by a continuous layer of T2 (due to the non-congruent melting character of the latter phase) with orientation relationships between these 'cousin' phases⁵. We have made a major advance in this field, and detail the spectacular morphological features associated with it.

After careful weighing of very high purity aluminium, lithium and copper base metals, an alloy with chemical composition $Al_{6.4}Cu_{0.9}Li_{2.7}$ was cast at 690 °C under a controlled atmosphere in a graphite mould. The alloy was cooled slowly from 620 to 570 °C with a solidification time of 1 h, as recorded by thermal analysis. Following casting and slow cooling to room temperature, the ingot was broken open. It exhibited well defined primary dendrites along free internal surfaces, these dendrites consisting entirely of large T₂ quasicrystals to the exclusion of the other equilibrium phases.

The macroscopic solidification morphology of the T2 dendrites is of major interest: they clearly consist of stacked triacontahedra (the triacontahedron is a non-regular polyhedron with 30 rhombus-shaped faces, 32 vertices, 60 edges and icosahedral

symmetry). These T2 dendrites, ~0.5 mm in diameter, can be seen in the low magnification scanning electron microscope images (Fig. 1). Their 2-, 3- and 5-fold symmetry axes, oriented perpendicularly to the picture plane, are shown in scanning electron micrographs in Fig. 2*a*, *b* and *c* respectively. They exhibit nearly perfect triacontahedral morphology (Fig. 2*d*), with the exception of a number of growth steps along some rhombic faces.

Such a triacontahedral morphology is a novel observation in crystallography and mineralogy. The triacontahedron is known to be the basic volume of the three-dimensional Penrose tiling^{6,7}. This morphology can undoubtedly be attributed to the quasicrystalline nature of the primary dendrites. X-ray powder diffraction and selected area electron diffraction patterns obtained from crushed macroscopic triacontahedra (Fig. 3) reveal that the primary dendrites consist solely of $T2-Al_6CuLi_3$ (with no occurrence of $R-Al_5CuLi_3$ or other phases such as aluminium).

The striking similarity between the morphological features of these packed triacontahedra (macroscopic scale) and the atomic structure model for quasicrystals (ref. 8) recently proposed is also of great interest. This atomic structure model involves triacontahedral atomic decoration of the two rhombohedral tiles of the 3-dimensional Penrose tiling^{6,7} (Fig. 4). On the other hand, the cubic $Im\bar{3}$ structure of the R phase close to T2 (refs 9, 10) consists of a regular packing of triacontahedra⁵.

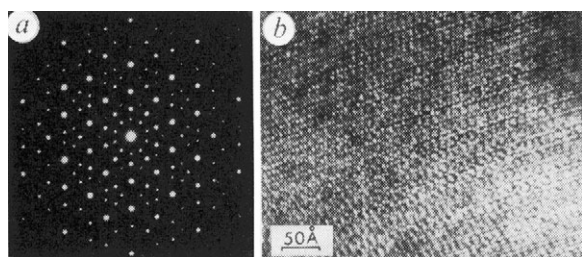


Fig. 3 *a*, Electron diffraction pattern and *b*, corresponding high resolution transmission electron microscopic image obtained along the five-fold axis from a fragment of crushed $T2-Al_6CuLi_3$ macroscopic triacontahedron.

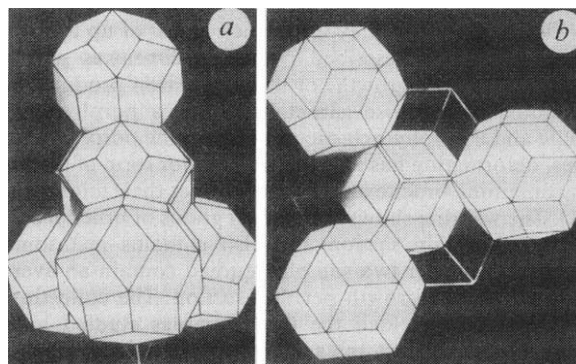


Fig. 4 Triacontahedra decorating the prolate (*a*) and oblate (*b*) rhombohedra of 3-dimensional Penrose tiling^{5,8}.

Using such samples, further work is now needed to describe more precisely the microscopic and geometrical features of this new morphology and to implement the structural models for quasicrystals. We propose that the results presented here on the icosahedral AlCuLi phase will lead to a better understanding and successful evaluation of the properties of this new type of material.

We thank R. Chamard, C. Bernard, P. Duroux, E. Janot and G. Duverneuil for their technical support and A. Dubus and F. X. Gilles for their contribution to previous steps of this study. Financial support of the Direction des Recherches Etudes et Techniques is gratefully acknowledged.

Received 1 August; accepted 18 September 1986.

1. Shechtman, D., Blech, I., Gratias, D. & Cahn, J. W. *Phys. Rev. Lett.* **53**, 1951–1953 (1984).
2. Sainfort, P., Dubost, B. & Dubus, A. *C.r. hebdom. Séanc. Acad. Sci. Ser. II, Paris* **301**, 689–692 (1985).
3. Hardy, H. K. & Silcock, J. M. *J. Inst. Metall.* **24**, 423–428 (1955).
4. Sainfort, P. & Dubost, B. in *Workshop on Aperiodic Crystals* (eds Michel, L. & Gratias, D.) *J. de Phys. Paris, Coll. C3* **47**, 321–330 (1986).
5. Audier, M., Sainfort, P. & Dubost, B. *Phil. Mag.* (in the press).
6. Duneau, M. & Katz, A. *Phys. Rev. Lett.* **54**, 2688–2691 (1985).
7. Elser, V. *Phys. Rev. B* **32**, 4892–4898 (1985).
8. Audier, M. & Guyot, P. in *Workshop on Aperiodic Crystals* (eds Michel, L. & Gratias, D.) *J. de Phys. Paris* **47**, Col. C3 **47**, 405–414 (1986).
9. Bergman, G., Waugh, J. L. T. & Pauling, L. *Acta crystallogr.* **10**, 254–259 (1957).
10. Cherkashin, E. E., Kripyakevich, P. I. & Oleksiv, G. I. *Sov. phys. Crystallogr.* **8**, 681–685 (1964).

Evidence for carbonate in the mantle

G. W. Berg

Geology Department, University of Port Elizabeth, PO Box 1600, Port Elizabeth 6000, South Africa

The presence of primary carbonate in kimberlite magmas, which erupt to the Earth's surface from depths of ≥ 150 km, carrying with them the rare elemental form of carbon, diamond, is evidence for a reservoir of carbon in the Earth's mantle. Hypotheses about the form in which this carbon is stored were largely speculative until laboratory experiments^{1,2} established that at depths of ≥ 50 km carbon can be incorporated in the solid carbonates, dolomite ($\text{CaMg}(\text{CO}_3)_2$) and magnesite (MgCO_3). However, only one instance of minute inclusions of dolomite in a sample of rock from the mantle has ever been recorded³, despite 25 years of intensive study of these rocks. Here I report the presence of hitherto undescribed brucite/calcite intergrowths ($\text{Mg}(\text{OH})_2/\text{CaCO}_3$) in mantle-derived xenoliths which occur in the diamond-bearing volcanic kimberlite pipes of Kimberley, South Africa. These intergrowths could reflect the presence of relatively abundant dolomite in the mantle. They seem to have been overlooked to date because of their fine textures and susceptibility to replacement by secondary serpentine.

The brucite and calcite are finely intergrown on a 10–20- μm scale (Fig. 1). Together, they form black grains of up to ~ 4 mm which are sometimes described in hand specimens as 'graphite' or 'carbon'. The intergrowths or their alteration products have been studied in a coarse garnet lherzolite, a porphyroclastic lherzolite and a porphyroclastic wehrlite. In the porphyroclastic samples, deformation has affected the exterior form of the black grains containing brucite and calcite, but not their microtexture (Fig. 1). The wehrlite shows about eight grains of brucite/calcite intergrowth per thin section and also contains metasomatic phlogopite. The other two samples studied contain an average of one grain of intergrowth per thin section. The collection of 70 nodules which yielded the three samples studied here is probably biased against metasomatized specimens. Thus the intergrowths may be more widely distributed in peridotite nodules from the Kimberley area than the present sampling indicates.

The brucite/calcite intergrowths are interpreted as reflecting the decarbonation products of dolomite, by analogy with the general reaction



used to describe the formation of pectatite and predazzite marbles, the periclase being hydrated to brucite⁴. The reaction in peridotite xenoliths is thought to have proceeded in response to decompression resulting from the rapid emplacement of their kimberlite hosts from depths of > 100 km. The exact path of the reaction is not certain. It is known to be complicated by solid-solution effects⁵ and the mantle analogue of dolomite in the equation is likely to be a solid solution between dolomite and magnesite⁶. The term dolomite is used here in this broad sense.

According to experimental data recently reviewed in ref. 7, a decrease in pressure to ~ 20 kbar could result in the reaction of mantle dolomite with coexisting silicates, with the accompanying evolution of CO_2 vapour. Eggler⁸ has suggested that these reactions may prevent dolomite from reaching the Earth's surface. However, the observations reported here suggest that the silicate-carbonate reactions are too slow to consume the dolomite, so that the metastable silicate-carbonate assemblages persist to ≤ 5 kbar, when dolomite decarbonates⁵ independently of surrounding silicates.

A wide range of oxygen fugacities has been postulated for the upper mantle⁹. The lower ranges (iron-wüstite buffer) are incompatible with the presence of carbonate. The dolomite inferred here is consistent with the EMOG (enstatite-magnesite-olivine-graphite) buffer¹⁰.

The preservation of the texture illustrated in Fig. 1 is probably dependent on the serpentinization of the host peridotite nodule being predominantly isochemical. An abundance of circulating water would remove both the brucite formed during the serpentinization of olivine in the nodule¹¹, and that thought to have formed from the decarbonation of dolomite. Fairly widespread serpentinization and removal of relatively soluble brucite from assemblages of the type illustrated in Fig. 1 may explain why the occurrence of these brucite/calcite intergrowths has not been reported before.

The dolomite inferred here to be reflected in the brucite/calcite assemblages probably represents carbonate from a shallower level in the mantle than the source of the kimberlite that brought the host peridotite nodule to the surface¹². The dolomite may be a product of mantle metasomatism, and whether such meta-

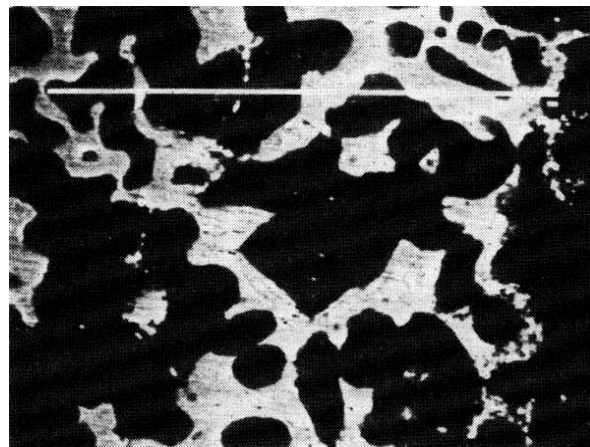


Fig. 1 Electron backscatter image of brucite/calcite intergrowth texture. Calcite (white) and brucite (black) were identified by microprobe analysis on polished thin sections and by X-ray diffraction of separated brucite/calcite grains. An X-ray raster image (not shown) shows traces of SiO_2 in the brucite which are thought to reflect incipient serpentinization. Serpentine was identified in more heavily altered brucite/calcite intergrowths by microprobe analysis. Scale bar, 100 μm .

somatism ultimately reflects plate tectonic processes or merely the interaction of magmas with wall-rock is subject to debate. However, it is suggested that the existence of the brucite/calcite assemblages described here resolves the contradiction of compelling experimental evidence for the presence of carbonate in the mantle with its apparent paucity in mantle xenoliths.

I thank R. S. Rickard and R. Moore for help with analyses on the Cameca microprobe of the Geochemistry Department at the University of Cape Town and the University of Port Elizabeth Research Committee for financial support.

Received 8 April; accepted 13 August 1986.

1. Wyllie, P. J. *J. Geol.* **86**, 687-713 (1978).
2. Eggler, D. H. *Am. J. Sci.* **278**, 305-343 (1978).
3. McGetchin, T. R. & Besançon, J. R. *Earth planet. Sci. Lett.* **18**, 408-410 (1973).
4. Harker, A. *Metamorphism* (Chapman and Hall, London, 1974).
5. Goldsmith, J. R. *J. geophys. Res.* **85**, 6949-6954 (1980).
6. Irving, A. J. & Wyllie, P. J. *Geochim. cosmochim. Acta* **39**, 35-53 (1975).
7. Wyllie, P. J., Huang, W. L., Otto, J. & Byrnes, A. P. *Tectonophysics* **100**, 359-388 (1983).
8. Eggler, D. H. *Yb. Carnegie Instn Wash.* **74**, 468-474 (1975).
9. Arculus, R. J. *Am. Rev. Earth planet. Sci.* **13**, 75-95 (1985).
10. Eggler, D. H. *Geophys. Res. Lett.* **10**, 365-368 (1983).
11. Hemley, J. J., Montoya, J. W., Shaw, D. R. & Luce, R. W. *Am. J. Sci.* **277**, 353-383 (1977).
12. Wyllie, P. J. *J. geophys. Res.* **85**, 6902-6910 (1980).

Lahars initiated by the 13 November 1985 eruption of Nevado del Ruiz, Colombia

Donald R. Lowe*, Stanley N. Williams*, Henry Leigh*,
Charles B. Connor†, J. Bruce Gemmell†
& Richard E. Stoiber†

* Department of Geology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

† Department of Earth Sciences, Dartmouth College, Hanover, New Hampshire 03755, USA

The eruption of Nevado del Ruiz in Colombia on 13 November 1985 was accompanied by the formation of four lahars (mud flows) triggered by the melting of glacial ice near the summit of the volcano. The lahars began as flows of water, sand and gravel, but they incorporated clay by eroding the soil along the steep valleys through which they passed. The largest flow was a cohesive debris flow, more than 45 m deep and moving at $\sim 12 \text{ m s}^{-1}$ when it debouched from the canyon of the Rio Lagunilla 2.5 km from Armero, where it killed 25,000 people. The continuing volcanic activity and abundant remaining glacial ice create an extremely high risk of future destructive flows.

At 5,400 m, Nevado del Ruiz is the highest of a cluster of five volcanic peaks in the Cordillera Central of north-central Colombia (Fig. 1) and is capped by a glacier 2-5 km wide. A major eruption of Ruiz occurred in 1595, and in 1845, Ruiz was the source of a lahar which killed more than 1,000 people at Ambalema (Fig. 1) on the Magdalena River (refs 1-3 and J. B. Tomblin, unpublished work). The recent activity of Ruiz began with earthquakes in November 1984, followed by a small phreatic eruption on 22 December (refs 4, 5, 6 and J. B. Tomblin, unpublished work). A second eruption occurred on 11 September 1985 (ref. 7).

On 13 November 1985, a minor eruption at $\sim 15:00$ local time produced a light ashfall north-east of the volcano⁸. Between 21:00 and 21:15 a larger eruption occurred, producing small pyroclastic flows and surges near the summit^{9,10} and melting a small part of the summit glacier, which generated four major lahars⁸ (Fig. 1). The two most destructive lahars moved eastwards down the Rio Lagunilla and its tributary, the Rio Azufrado, toward Armero; a third moved northeastwards down

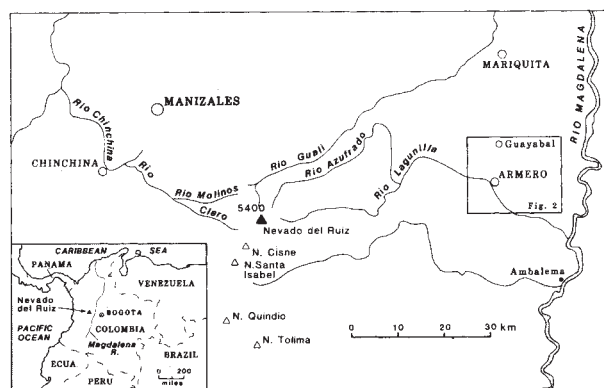


Fig. 1 Generalized location map of Nevado del Ruiz and vicinity, and (inset) Colombia.

the Rio Guali toward Mariquita; and a fourth moved northwards down the Rio Molinos into the Rio Claro and thence to the Rio Chichina toward Chichina (Fig. 1).

Examination of the Rio Guali lahar deposits on the north-east slope of Ruiz, ~ 5 km from the summit, indicates that the initial flow was a mixture of sand, gravel, water and probably ice, but contained little clay. The main effects of the flow at this point were to strip away a 0.5-2-m-thick layer of vegetation, soil and older pumice overlying massive volcanic rock and to deposit units of coarse, moderately-to-poorly sorted, well-stratified gravel and sand. Most of the clay in the lahars that emerged from the canyons below was evidently derived by erosion of soils along the steep, V-shaped valleys through which the flows passed.

The valleys on the east and north-east side of Ruiz descend from elevations of $\sim 5,300$ m to only 400 m, where they enter the Magdalena River valley over river distances of 60-80 km. Headward slopes are generally $10-15^\circ$; slopes near the base are generally $3-4^\circ$. The lahars stripped vegetation and soil from the canyon walls to heights commonly exceeding 50 m and eroded many older alluvial deposits from the valley bottoms. The Rio Lagunilla flow travelled ~ 60 km before debouching just west of Armero. Based on a 2-h travel time (initiation at $\sim 21:30$ to passage through Armero at $23:30$), the estimated mean flow velocity was 8.3 m s^{-1} . The course of the Rio Azufrado flow was 12 km longer than that down the Rio Lagunilla, giving a mean velocity of 10.0 m s^{-1} . Examination of the river valleys, as well as eye-witness reports, indicate that the Rio Azufrado flow was considerably more voluminous than the flow in the Rio Lagunilla, and that the former carried the bulk of the rocky debris. The differences in path-length and composition of these flows suggest that, downstream from the junction of the two rivers, the flows may have passed as a series of two or perhaps more discrete surges. Cross-cutting relationships at the junction of the Rio Lagunilla and Rio Azufrado indicate that the Azufrado flow reached the junction first¹¹, and may have formed the first surge to inundate Armero. The passage of a small lahar in the Azufrado in September 1985 (ref. 12) may have facilitated the movement of the later lahar.

About 1 km upstream from the canyon mouth, 2.5 km west of Armero, the Rio Lagunilla flows followed a sharp bend to the right, banking high onto the left (north) wall of the canyon (Fig. 2). The mean velocity (w) of debris flows at maximum discharge can be calculated indirectly from their superelevation around bends:

$$w = [g \cos \beta \tan \delta R]^{1/2} \quad (1)$$

where g is the acceleration of gravity, β is the slope of the stream bed, δ is the slope of the banked flow surface (here measured from differential levels of scour on opposite sides of the river) and R is the radius of curvature of the bend^{13,14}.

Around this bend, $g = 9.8 \text{ m s}^{-2}$, $\beta = 4^\circ$, $\delta = 7.5^\circ$ and $R = 110 \text{ m}$, yielding a mean flow velocity of 11.9 m s^{-1} . The flow was $\sim 45 \text{ m}$ deep at mid-channel and had a cross-sectional area of nearly $4,000 \text{ m}^2$. The calculated discharge (Q) was $\sim 47,500 \text{ m}^3 \text{ s}^{-1}$. Similar flows at Mt St Helens in the United States in 1980 had peak velocities of $3.2\text{--}40 \text{ m s}^{-1}$ and peak discharges of $2,440$ to $>190,000 \text{ m}^3 \text{ s}^{-1}$ (ref. 15). Most lahars and debris flows attain their maximum depths near the flow fronts¹³. The conditions calculated above probably characterized the first surge of debris to leave the Rio Lagunilla canyon.

The difference between estimated mean flow velocities and that of the initial surge immediately upstream from Armero may reflect imprecise estimates of the times of flow initiation and arrival; down-slope changes in channel geometry and resistance; and down-slope flow evolution from friction-dominated gravel-sand-water slurries to cohesive gravel-sand-clay-water debris flows.

When the lahars left the canyon of the Rio Lagunilla, they emptied onto the gently sloping surface of the west side of the Magdalena River valley (Fig. 3). Within 2 km of the canyon mouth, the Rio Lagunilla is incised 10–15 m below this surface. The initial flow, however, was sufficiently deep that it spread radially from the canyon mouth and split into three main lobes. The northern lobe moved to the ENE, left the Rio Lagunilla, swept through northern Armero following a broad distributary valley, and then turned northward and flowed as a series of anastomosing lobes to within 1.5 km of Guayabal (Fig. 3).

Most of the debris flowed southeastwards, following the broad valley of the Rio Lagunilla. Because the present river lies along the southern edge of this valley, and was too small to accommodate the volume of the lahar, most of the southern lobe moved north of the river, inundating and obliterating the low-lying southern half of Armero. The only parts of the city not completely destroyed were the slightly elevated central and eastern areas between the two main flow lobes. The inertia of the flow carried all but a small part of the debris away from the present river to the south-east and east for nearly 16 km, along a broad alluvial valley which may be an abandoned river course or a route of previous lahars (Fig. 3). A much smaller amount of debris followed the course of the Lagunilla, forming a small third lahar lobe.

Early reports from Armero indicated that many victims were burned, suggesting the possibility that the lahars were hot. Although virtually everyone caught in the flow suffered extensive lacerations because of the abrasive power of the coarse debris, we saw no evidence that the flows were hot. One victim interviewed several weeks after the disaster states that the flows were cold but still burned. The high sulphate content of the flow deposits¹⁶ suggests that the 'burning' was actually caused by the acidity of the flow.

The Lagunillas-Azufrado lahar deposits are generally between 1 and 2 m thick and mantle, rather than fill, pre-flow topography. They cover an area of nearly 32 km^2 , which, for an average depth of 1.5 m, implies a volume of $\sim 4.8 \times 10^7 \text{ m}^3$. The deposits consist of grey, dense, massive, extremely poorly sorted clayey gravelly sand. Granulometric analysis of one sample from the upper part of the proximal flow deposits 0.5 km south-east of the canyon mouth shows (by weight) 47.9% sand, 27% gravel, 17.5% silt and 7.6% clay. All the coarser debris is composed of lithic volcanic rock types; fresh pumice was not identified. Clasts $\geq 50 \text{ cm}$ in diameter are generally rounded and probably eroded from older alluvial deposits. Clasts $< 15 \text{ cm}$ in diameter are commonly matrix-supported.

The grain-size structuring of the deposits indicates that, within the debris flows, the mixture of fluid and granular materials $\leq 10 \text{ cm}$ in diameter behaved as a homogeneous phase in which there was no separation or differential settling of constituents. Clasts $\geq 15 \text{ cm}$ across, locally reaching 10 m in diameter, occur only at the base of the lahar deposits. They were either transported at the bases of the flows or settled as the flows decelerated.



Fig. 2 Photograph of the Rio Lagunilla canyon looking upstream, $\sim 2.5 \text{ km}$ from Armero and 1.5 km upstream from the mountain front. Lahars generated by the 13 November 1985 eruption of Nevado del Ruiz have stripped vegetation and soil from the lower parts of the canyon walls and deposited a thin veneer of mud. A man (arrowed) gives the scale. The superelevation of the flow along the north bank (the right bank in the photograph) can be seen.

As a result, many of the deposits show normal grading of the coarsest constituents.

The observations that the flow deposits mantle, rather than fill, the topography and show differential settling of larger clasts suggest that the strength of the flow was relatively low. Flow strength (k) can be estimated by¹³:

$$k = T\rho_f g \sin \alpha \quad (2)$$

where T is the deposit thickness, ρ_f is the flow density and α is the slope of the deposit surface¹³. Over broad areas, $T \approx 100\text{--}150 \text{ cm}$ and $\alpha < 1.4^\circ$, and ρ_f is taken as 2.2 g cm^{-3} . For these deposits, $k \approx (5.3\text{--}8.0) \times 10^3 \text{ dyn cm}^{-2}$, which is comparable to that of many other debris flows^{13,17}.

The large volume of glacial ice remaining on Nevado del Ruiz after 13 November, the abundance of debris around the summit, and the precipitous, narrow gorges leading downward to river valleys around the volcano collectively represent ideal conditions for the formation and long-distance movement of lahars.

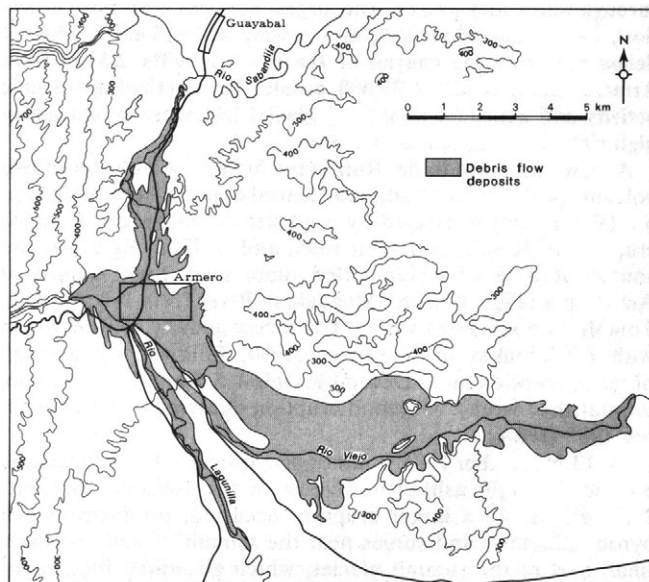


Fig. 3 Map of the Armero region (box in Fig. 1), showing the distribution of the deposits of the Rio Lagunilla lahar.

It is clear that the volcano is presently active and the probability of future eruptions is high. Even minor volcanic activity would inevitably trigger new lahars. Appropriate measures for early detection and hazard mitigation should be taken.

Supported by NSF (grant EAR86-04158 to S.N.W.), the National Geographic Society Magazine (grant to S.N.W.), the Department of Geology, Louisiana State University and the Department of Earth Sciences, Dartmouth College (Stoiber Field Fund grant to J.B.G. and C.B.C.). We thank the Fuerza Aereo Colombiana for logistical support and especially Brigadier General Melendez (FAC), NBC and ABC television, the US Army, Mr Paul Bell (US State Department), the Universidad Nacional (Manizales, Colombia), the US Geological Survey, and the following colleagues for assistance; Roberto Leigh, Omar Pedraza, Antonio Rivera, William Fritz, Hobart McDowell, Steve Raymer and Eric Johnson.

Received 14 April; accepted 13 August 1986.

- Hantke, G. & Parodi, A. *Catalog of the Active Volcanoes of the World* Pt XIX, 2-3 (International Association of Volcanology, Rome, 1959).
- Herd, D. G. *Glacial and Volcanic Geology of the Ruiz-Tolima Volcanic Complex, Cordillera Central, Colombia* 11 (INGEOMINAS, Bogotá, 1982).
- Simkin, T. *et al. Volcanoes of the World* 95 (Hutchinson Ross, Stroudsburg, 1981).
- Hall, M. L., Tomblin, J. & Gómez, O. *SEAN Bull.* 10(5), 5 (1985).
- Salazar-A., B., Calvache-V., M. L. & García-P., N. *SEAN Bull.* 10(7), 10 (1985).
- Salazar-A., B. & Mejia, G. *SEAN Bull.* 10(8), 8 (1985).
- Jaramillo, L. *et al. SEAN Bull.* 10(9), 3 (1985).
- Herd, D. G. *et al. Eos* 67, 457-460 (1986).
- Calvache-V., M. L. *EOS* 67, 405 (1986).
- Janda, R. J., Banks, N. G., Pierson, T. C., Calvache-V., M. L. & Thouret, J. C. *Eos* 67, 406 (1986).
- Pierson, T. C., Janda, R. J., Thouret, J. C. & Calvache-V., M. L. *EOS* 67, 406 (1986).
- Cepeda, H. *et al. EOS* 67, 407 (1986).
- Johnson, A. R. *Physical Processes in Geology* 433-534 (Freeman, San Francisco, 1979).
- Johnson, A. R. & Rodine, J. R. in *Slope Instability* (eds Brunsden, D. & Prior, D. B.) 257-362 (Wiley, Chichester, 1984).
- Pierson, T. C. *Geol. Soc. Am. Bull.* 96, 1056-1069 (1985).
- Williams, S. N. *et al. Science* 233, 964-967 (1986).
- Johnson, A. R. & Rodine, J. R. in *Slope Instability* (eds Brunsden, D. & Prior, D. B.) 294 (Wiley, Chichester, 1984).

Sporadic shutdown of North Atlantic deep water production during the Glacial-Holocene transition?

W. H. Berger* & E. Vincent†

* Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA

† Scripps Institution of Oceanography and Département de Géologie Dynamique, Université Pierre et Marie Curie, 75005 Paris, France

Pulsations in the production of North Atlantic deep water (NADW) have been implicated in generating drastic climatic fluctuations during the Glacial-Holocene (G/H) transition¹⁻³. The stable isotope record of benthic foraminifera in high-resolution cores from the Norwegian Sea suggests that such pulsations did occur⁴. Although the question of exact timing (and mechanism) is still open there is little doubt that NADW pulsations were important in climatic history because the rate of NADW production influences the rate of advection of heat to the northern North Atlantic⁵. Here we report that a sporadic shutdown of NADW may be recognizable in deep-sea carbonates with normal (low) sedimentation rates. Hence the possibility arises that relatively short-lived events (~1,000-2,000 yr) in deep circulation can be mapped over large areas of the sea floor, despite the detrimental effects of bioturbation on signal resolution.

The overall pattern of deep circulation in the present ocean is characterized by the asymmetry in deep-water properties of the Pacific and the Atlantic, which reflects the difference in age between sub-thermocline waters and, in essence, results from NADW production. We have chosen four cores to recon-

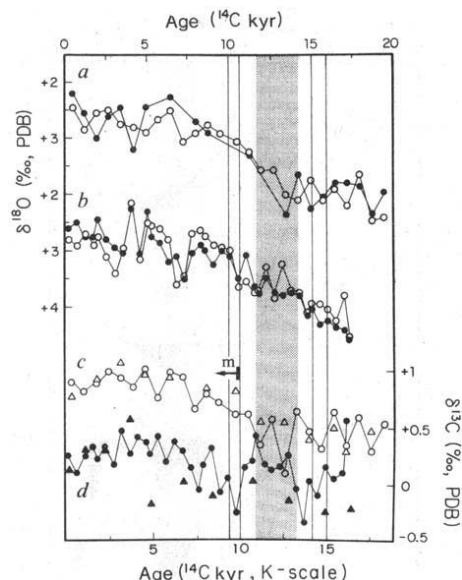


Fig. 1 Oxygen and carbon isotopes in the shells of benthic foraminifera from deep-sea sediments in the central Atlantic and western equatorial Pacific. Age scales are based on radiocarbon datings of bulk sediment (top-scale, refs 7, 8) and on correlation of planktonic oxygen isotope signals with the Norden (radiocarbon-) timescale (bottom scale, ref. 18). *a*, $\delta^{18}\text{O}$ values for *Cibicidoides wuellerstorfi* (○) and *Oridorsalis umbonatus*, adjusted by -0.6% (●), in core INMD 115 Bx, Atlantic. *b*, $\delta^{18}\text{O}$ values for *C. wuellerstorfi* (○) and *O. umbonatus*, adjusted by -0.75% (●) in core ERDC 112 Bx, Pacific. *c*, $\delta^{13}\text{C}$ values for *C. wuellerstorfi* in cores INMD 115 Bx (○) and INMD 113 Bx (△), Atlantic; m over arrow, appearance of *Globorotalia menardii*. *d*, $\delta^{13}\text{C}$ values for *C. wuellerstorfi* in cores ERDC 112 Bx (●) and ERDC 123 Bx (▲), Pacific. Atlantic data from Table 1, Pacific data from ref. 6.

struct the sequence of deep-water age difference in the two oceans since maximum glaciation (Fig. 1): ERDC 112 Bx, ERDC 123 Bx, INMD 113 Bx, and INMD 115 Bx. The ERDC box cores are from the Ontong Java Plateau in the western equatorial Pacific, from depths of 2,169 m and 2,948 m, respectively. Isotope stratigraphies and details on sedimentation rates and physical properties have been given elsewhere^{6,7}. The INMD box cores are from the Mid-Atlantic Ridge, at 15.5 and 17.4° S, and from depths of 3,471 m and 3,427 m, respectively. Core data are given in ref. 8 and isotope data are listed in Table 1. Core ERDC 112 Bx and the INMD cores are close to the present two-degree potential temperature level⁸. The INMD sites are within the high-salinity-core layer of (lower) NADW. The ERDC sites sample general Pacific deep water with average salinity.

The benthic foraminifera analysed are *Cibicidoides wuellerstorfi* and *Oridorsalis umbonatus*. Both species were used for plotting $\delta^{18}\text{O}$ values, but only *C. wuellerstorfi* was used for the carbon isotope record. We believe^{10,11}, as do others¹²⁻¹⁶, that this species best reflects the $\delta^{13}\text{C}$ content of deep waters, with a minimum of interference from vital effects and contamination by interstitial values of $\delta^{13}\text{C}$. Analyses were carried out as described in ref. 17.

The oxygen isotope record in the deep Atlantic (INMD 115 Bx, Fig. 1a) shows the rise from heavy to light oxygen isotope values during the G/H transition, which is largely due to the addition of meltwater. The change occurs mostly between 14,000 and 8,000 yr BP; the G/H range is between 1.3 and 1.4‰. Of the two foraminifera, *C. wuellerstorfi* is plotted as measured, while the values for *O. umbonatus* are adjusted by -0.6% to make them congruent. Exact ages cannot be assigned because of the effects of resedimentation and bioturbation on radiocarbon dates; a measure of uncertainty is

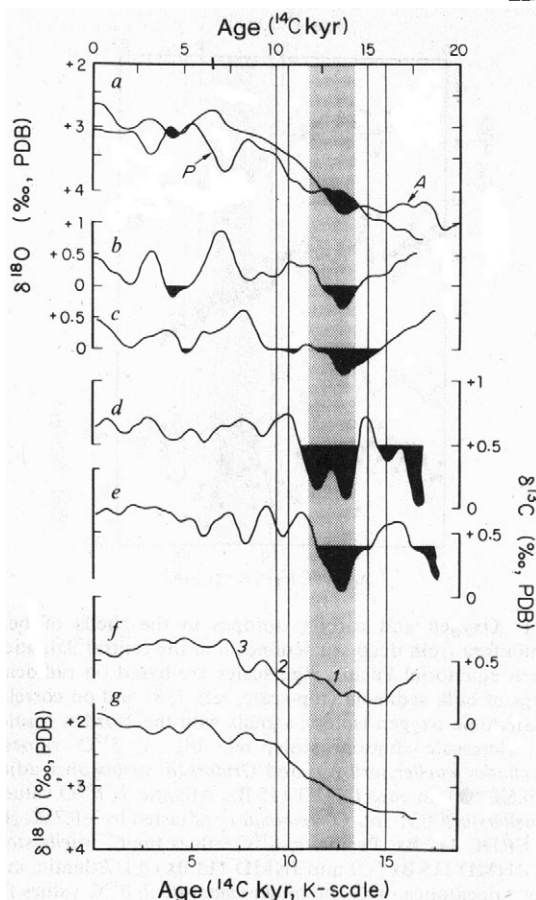


Fig. 2 Smoothed curved (triangle filter, 12321) derived from data in Fig. 1, *C. wuellerstorfi* in INMD 115 Bx (Atlantic) and in ERDC 112 Bx (Pacific). *a*, $\delta^{18}\text{O}$ values in Atlantic (A) and Pacific (P); black areas: P-values lighter than A-values. *b*, Difference in $\delta^{18}\text{O}$ values Atlantic–Pacific; black: difference negative. *c*, As *b*, after shifting Pacific curve to right by 1,000 yr. *d*, Difference in $\delta^{13}\text{C}$ values Atlantic–Pacific; black: difference $< +0.5\%$. *e*, As *d*, after shifting Pacific curve to right by 1,000 yr. *f*, Average of Atlantic and Pacific $\delta^{13}\text{C}$ curve. *g*, Average of Atlantic and Pacific $\delta^{18}\text{O}$ curve.

given by the alternative timescales shown. The timescale on top is the best-guess scale derived from bulk radiocarbon ages as listed in ref. 8. The timescale on the bottom¹⁸ is derived from a correlation of the oxygen-isotope signal with the climate-stratigraphy of the Norden scale¹⁹. The stippled pattern denotes the timespan for which sporadic shutdown of NADW is proposed.

A similar oxygen isotope plot for the Pacific (ERDC 112 Bx, Fig. 1*b*) shows a noisier record (possibly because of lumpy bioturbation in this more productive area) but otherwise reflects the same transitional change. *Oridorsalis umbonatus* in this record was adjusted by -0.75% . The greater difference between the two foraminiferal species may be a function of the higher productivity in the Pacific. The G/H range, here too, is $\sim 1.4\%$. If a range of 1.2% is accepted for the ice-effect²⁰ then the remaining 0.2% would suggest that glacial deep waters were roughly 1°C colder than postglacial ones, in both oceans. The main conclusion from the oxygen isotope records is that fully glacial conditions were reached at the bottom of each of the box cores, and that the transition is well represented, apparently without hiatus.

The carbon isotope records show the Pacific/Atlantic asymmetry referred to above, with lighter $\delta^{13}\text{C}$ values in the deep Pacific, which reflects the older age of the water masses there (Fig. 1*c,d*). For the past several thousand years the difference

in $\delta^{13}\text{C}$ is close to 0.6% , as expected from direct measurement of water samples in the present ocean²¹. This corresponds to an age difference of ~ 600 yr (0.1% in $\delta^{13}\text{C}$ being equivalent to ~ 100 yr (ref. 11)). The difference persists back to $\sim 10,000$ yr BP, suggesting that there has been no major reorganization of deep circulation within this timespan. The records confirm the diminished importance of Atlantic deep-water production during glacial time, an hypothesis put forward previously on the basis of heat budget considerations²² and benthic isotope data similar to ours^{23–26}.

Within a narrow time interval during deglaciation proper (the shaded zones in Figs 1 and 2) differences in $\delta^{13}\text{C}$ values sometimes disappear. We propose that this is the result of a shutdown of NADW. The decrease in Atlantic/Pacific $\delta^{13}\text{C}$ contrast during deglaciation owes as much to anomalously heavy values in the Pacific foraminifera as it does to anomalously light values in the Atlantic. It is conceivable therefore that an effect besides NADW production has to be considered in this context, namely North Pacific deep water production (see ref. 27). It is difficult to explain these anomalies as the result of bioturbation, because first, the δ -values in question are at the limit of the total range (how could mixing produce extremes?) and second, anomalous values in both Atlantic and Pacific coincide although correlation between the cores is independent of the $\delta^{13}\text{C}$ -stratigraphies.

The point at which NADW production sets in fully and permanently, at the beginning of the Holocene, is marked by the reappearance in the Atlantic of the previously regionally-extinct planktonic foraminiferan *Globorotalia menardii* (Fig. 1*c*). The coincidence between the two events may be more than fortuitous. The sinking of NADW in the Norwegian Sea implies transport of surface water across the Equator. According to Gordon²⁸ this water is mainly transported around the Cape of Good Hope, from the Indian Ocean, after southward flow through the Mozambique Channel. This southward flow carries tropical fauna. Thus, the beginning of NADW production favours the reseeding of the tropical foram *G. menardii* in the Atlantic, through the entrainment of tropical Indian Ocean fauna by a strengthened Agulhas Current. During glacial conditions the southward flowing current in the Indian Ocean was weakened, and colder^{29,30} and was not effective in conveying *G. menardii* to the Cape. Also, entrainment around the Cape was diminished.

The $\delta^{13}\text{C}$ values show an overall Glacial to Holocene shift from light to heavy of between 0.4 and 0.5% (Fig. 1*c,d*). Such a shift is expected from the buildup of organic carbon in the terrestrial biosphere and in soils (Shackleton's 'rain forest' effect³¹). The biosphere build-up preferentially extracts isotopically-light carbon from the ocean/atmosphere system, leaving the ocean enriched in ^{13}C . A late Holocene shift towards lighter values is evident, which we ascribe to biomass destruction (desertification) following the climatic optimum.

As long as the NADW production takes place, the deep Atlantic should remain warmer than the deep Pacific⁹. If NADW production shuts off, the temperature in the deep Atlantic should fall and reach values more like those in the Pacific. To test this possibility we have superimposed the Pacific and Atlantic oxygen isotope data, after smoothing with a five-point triangle filter (Fig. 2*a*). For the past 10,000 yr the Pacific record shows the heavier (colder) values, although not the full difference expected, because of the shallower water depth of the sample cores in the Pacific (2,169 m versus 3,427 m in the Atlantic). The only time when this Atlantic/Pacific contrast reverses sign is in the early part of deglaciation, during the same interval in which the $\delta^{13}\text{C}$ difference tends to disappear sporadically. Thus, the hypothesis of NADW shutdown is supported by the oxygen isotope record.

The observed reversal of the oxygen isotope difference in the early stage of deglaciation is quite robust towards changes in timescale (that is, towards miscorrelation of the records)—see Fig. 2*b,c*. Figure 2*b* shows the record of the $\delta^{18}\text{O}$ difference seen

Table 1 $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ (‰, PDB) values in benthic foraminifera in INMD box cores from the South Atlantic

Depth in core (cm)	<i>C. wuellerstorfi</i>		<i>O. umbonatus</i>
	$\delta^{18}\text{O}$	$\delta^{13}\text{C}$	$\delta^{18}\text{O}$
INMD 115 Bx			
0-1	2.45	0.90	2.81
1-2	2.86	0.83	3.17
2-3	2.56	0.90	3.60
3-4	2.51	1.01	3.22
4-5	2.77	0.95	3.07
5-6	2.80	0.88	3.78
6-7	2.92	1.03	3.05
7-8	2.67	0.78	—
8-9	2.50	0.99	2.83
9-10	3.04	0.96	—
10-11	2.89	0.67	3.29
11-12	2.77	0.79	3.52
12-13	2.90	0.73	—
13-14	3.03	0.63	—
14-15	3.24	0.62	3.89
15-16	3.53	0.35	—
16-17	3.56	0.57	—
17-18	4.00	0.09	4.97
18-19	4.10	0.65	4.25
19-20	3.73	0.47	4.85
20-21	4.11	0.33	4.63
21-22	3.88	0.64	4.40
22-23	4.19	0.36	4.39
23-24	3.65	0.60	4.47
24-25	4.46	0.29	4.93
25-26	4.39	0.56	4.56
INMD 113 Bx			
0-1		0.77	
2-3		0.93	
4-5		1.08	
6-7		0.97	
8-9		0.95	
11-12		0.86	
13-14		0.82	
15-16		0.56	
17-18		0.56	
19-20		0.40	
21-22		0.50	
22-23		0.31	
24-25		0.47	

in Fig. 2a. If one assumes that the Pacific record should be shifted to the right by 1,000 yr to minimize the Pacific/Atlantic contrast, the reversal during early deglaciation still persists. A substantial (and unrealistic) shift of the Pacific record to the left, would be necessary to make this reversal disappear. The carbon isotope anomalies in the Pacific/Atlantic asymmetry are likewise quite robust towards miscorrelation (Fig. 2d,e).

When averaging the carbon isotope records (ERDC 112 Bx and INMD 115 Bx) an indication of stepwise change during deglaciation becomes apparent (steps labelled 1, 2, 3 on Fig. 2f). These steps may reflect abrupt changes in the deep-water carbon composition of the world ocean. It is interesting that they correspond to the major warming episodes proposed in the Norden scheme¹⁹. If these carbon shifts are due to storage of isotopically-light carbon in forests (and soils) one would then conclude that the deglacial growth of the biosphere proceeded in lockstep with the fluctuations in climate, each major warming episode producing a major increase in biosphere storage. The process would have started ~13,000 yr ago.

The average oxygen isotope record (Fig. 2g) shows a gradually accelerating change from the heavier glacial values to the Holocene ones, with a G/H range near 1.4‰. Substantial pre-13 kyr deglaciation^{32,33} is not evident in this benthic signal. Instead, these data support a late-deglaciation model^{18,34}. Such

a model strengthens the case for sporadic shutdown of NADW because a high rate of meltwater influx and of warming tends to prevent the sinking and downward mixing of surface waters in the North Atlantic, during deglaciation¹.

This research was supported by NSF grant OCE85-01578. We acknowledge discussions with J.-C. Duplessy, L. Labeyrie, and N. Shackleton, on carbon isotope stratigraphy of the late Pleistocene ocean. Analyses of the $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ values were carried out at Scripps Institution of Oceanography under the supervision of J. S. Killingley.

Received 10 March; accepted 22 August 1986.

- Olausson, E. *Progr. Oceanogr.* **3**, 221-252 (1965).
- Berger, W. H. & Killingley, J. S. *J. mar. Res.* **40**, 27-38 (1982).
- Broecker, W. S., Peteet, D. M. & Rind, D. *Nature* **315**, 21-26 (1985).
- Jansen, E. & Erlenkeuser, H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **49**, 189-206 (1985).
- Berger, W. H. & Labeyrie, L. (eds) *Abrupt climatic change—evidence and implications* (Reidel, Dordrecht, in the press).
- Vincent, E., Killingley, J. S. & Berger, W. H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **33**, 221-230 (1981).
- Berger, W. H. & Killingley, J. S. *Mar. Geol.* **45**, 93-125 (1982).
- Berger, W. H., Killingley, J. S., Metzler, C. V. & Vincent, E. *Quat. Res.* **23**, 258-271 (1985).
- Mantyla, A. W. & Reid, J. L. *Deep-Sea Res.* **30**, 805-833 (1983).
- Vincent, E., Killingley, J. S. & Berger, W. H. *Nature* **289**, 639-643 (1981).
- Berger, W. H. & Vincent, E. *Geol. Rdsch.* **75/1**, 249-269 (1986).
- Woodruff, F., Savin, S. M. & Douglas, R. G. *Mar. Micropaleont.* **5**, 3-11 (1980).
- Belanger, P. E., Curry, W. B. & Matthews, R. K. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **33**, 205-220 (1981).
- Graham, D. W., Corliss, B. H., Bender, M. L. & Keigwin, L. D. *Mar. Micropaleont.* **6**, 483-497 (1981).
- Duplessy, J.-C. *et al. Quat. Res.* **21**, 225-234 (1984).
- Corliss, B. *Nature* **314**, 435-438 (1985).
- Berger, W. H. & Killingley, J. S. *Science* **197**, 563-566 (1977).
- Berger, W. H., Killingley, J. S. & Vincent, E. *Nature* **314**, 156-158 (1985).
- Mangerud, J., Andersen, S. T., Berglund, B. E. & Bonner, J. J. *Boreas* **3**, 109-128 (1974).
- Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39-55 (1973).
- Kroopnick, P. M. *Deep-Sea Res.* **A32**, 57-84 (1985).
- Weyl, P. K. *Met. Monogr.* **8**, 37-62 (1968).
- Duplessy, J.-C., Chenouard, L. & Vila, F. *Science* **188**, 1208-1209 (1975).
- Streeter, S. S. & Shackleton, N. J. *Science* **203**, 168-170 (1979).
- Boyle, E. A. & Keigwin, L. D. *Science* **218**, 784-787 (1982).
- Shackleton, N. J., Imbrie, J. & Hall, M. A. *Earth planet. Sci. Lett.* **65**, 233-244 (1983).
- Shackleton, N. J. *Eos* **66**, 292 (1985).
- Gordon, A. L. *Science* **227**, 1030-1033 (1985).
- Vincent, E. *Mar. Biol.* **9**, 1-235 (1976).
- Prell, W. L. & Hutson, W. H. *Science* **206**, 454-456 (1979).
- Shackleton, N. in *The Fate of Fossil Fuel CO₂ in the Oceans*, 401-427 (Plenum, New York, 1977).
- Ruddiman, W. F. & McIntyre, A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 145-214 (1981).
- Ruddiman, W. F. & McIntyre, A. *Quat. Res.* **16**, 125-134 (1981).
- Mix, A. C. & Ruddiman, W. F. *Quat. Sci. Rev.* **4**, 59-108 (1985).

Giant subtidal stromatolites forming in normal salinity waters

Robert F. Dill*, Eugene A. Shinn†, Anthony T. Jones‡, Kevin Kelly‡ & Randolph P. Steinen§

* 610 Tarento Drive, San Diego, California 92106, USA

† Department of the Interior, Geological Survey, Fisher Island Station, Miami Beach, Florida 33139, USA

‡ University of Hawaii, Honolulu, Hawaii 96822

§ University of Connecticut, Storrs Connecticut 06268, USA

We report here the discovery of giant lithified subtidal columnar stromatolites (>2 m high) growing in 7-8 m of clear oceanic water in current-swept channels between the Exuma Islands on the eastern Bahama Bank. They grow by trapping ooid and pelletal carbonate sand and synsedimentary precipitation of carbonate cement within a field of giant megaripples. The discovery is important to geologists and biologists because similar organo-sedimentary structures built by a combination of cementation and the trapping of sediment by microbes were the dominant fossil types during the Precambrian. Stromatolites are thought to have been responsible for the production of free oxygen and thus the evolution of animal life^{1,2}. Until the discovery of small lithified subtidal columnar stromatolites in the Bahamas³, the only subtidal marine examples known to be living while undergoing lithification were

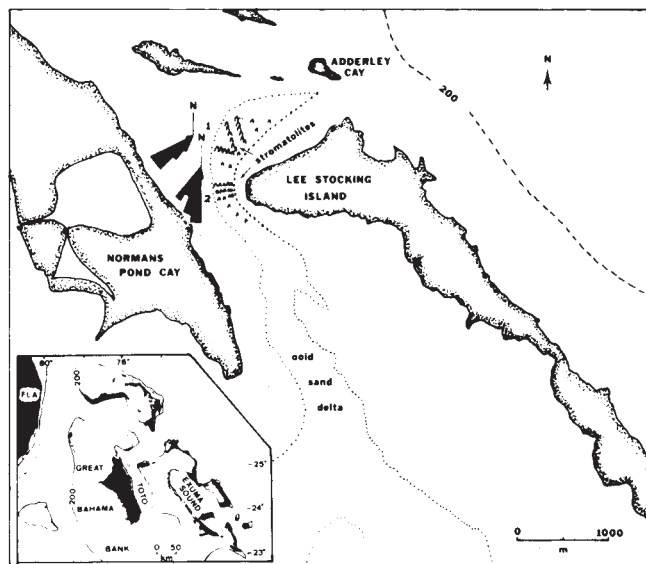


Fig. 1 Location of stromatolite-bearing tidal channel north and west of Lee Stocking Island, Exuma Cays, Bahamas. Rose diagrams are plots of the long axis of streamlined stromatolites that parallel the main axis of tidal flow in the channel: 1, 044–266° ($N=31$); 2, 014–194° ($N=60$). Orientation measurements of stromatolites were made with underwater Brunton at stations numbered 1 and 2. Visual observations demonstrate that many of the streamlined stromatolites are tilted toward the incoming tidal bottom currents. Contours in metres.

in the hypersaline waters of Hamelin Pool at Shark Bay, Western Australia^{4–7}. Shark Bay stromatolites range from intertidal to the shallow subtidal with the larger columns reaching 1 m in height. The Shark Bay stromatolites have strongly influenced geological interpretation; by analogy, many ancient stromatolites have been considered to have grown in intertidal and/or hypersaline conditions⁸, although hypersalinity was not a necessity for growth during the Precambrian because grazing metazoan life had not then evolved.

Stromatolites from the Exuma Islands occur in one of the many channels that exchange normal oceanic water from the deep Exuma Sound through the Pleistocene eolian islands to the shallows of the eastern Bahama Bank. Those described here are from the channel separating Lee Stocking Island from Adderley and Normans Pond Cays (Fig. 1). The Exuma Islands are well known to carbonate geologists for their spectacular eolian carbonate dunes and ooid tidal deltas^{9,10}. Exuma Sound has precipitous margins and a maximum depth of ~1.5 km. The eastern portion of Exuma Sound is open to the Atlantic Ocean and its waters which flood the stromatolites at high tide are exceptionally clear and of normal oceanic salinity (37–38‰). Underwater visibility in the tidal channel during flood tide is often >20 m. Ebb tidal waters from the shallow banks are more saline (as much as 40‰) due to evaporation, and have a visibility of 5–10 m, depending on weather conditions. Tides are diurnal and create currents between islands ranging between 60 and 100 cm s⁻¹.

To date, we have found growing stromatolites only associated with migrating 1- to 2.5-m-high ooid sand dunes. The stromatolites occur as individuals or as large, coalesced bioherms that form rows perpendicular to tidal flow (Fig. 2) and as randomly-spaced individuals of variable height. Randomly-spaced individuals are usually grouped in irregular-shaped patches 5–10 m across. The tops of the larger fence-like stromatolites are parallel with the crests of submarine sand dunes. Nearly all individuals, regardless of size, are streamlined by tidal currents. Many display higher growth rates on the side facing the incoming tide, which causes them to lean noticeably towards the clear water.

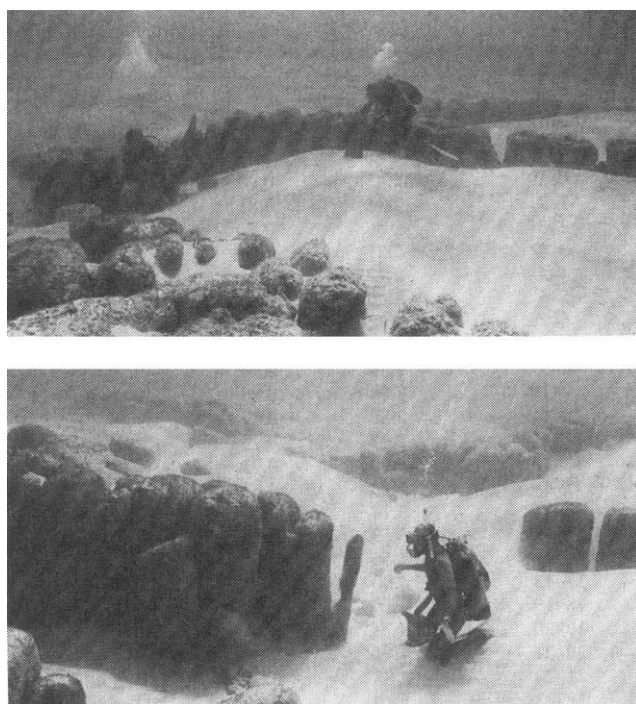


Fig. 2 Two views of large aligned fence-like stromatolites. Note that the height of migrating rippled oolitic sand dunes (megaripples) just equals the height of stromatolites and hence periodically covers most stromatolites. Smaller stromatolites in foreground of upper photograph are populated by the green alga *Batophora* spp. *Batophora* is uncommon or absent from actively growing specimens, such as in Fig. 3a and c, and those recently uncovered by migration of oolitic megaripples and bars. Photo taken at slack tide at a water depth of 7 m.

Many larger stromatolites have a soft botryoidal or pustular surface (Fig. 3a). Individual 1- to 10-cm-diameter botryoids are composed of uncemented ooid sand loosely bound by mucus and unidentified microbial filaments (Fig. 3b). Small stromatolites as much as 1 m high may also have a pustulate surface, but most are smooth and slightly elongate (streamlined) parallel to tidal currents. The surfaces of the small smooth individual are usually hard or crusty and lack the loosely held sand coating (Fig. 3c). Microbial filaments (Fig. 3b) occupy only the upper 2–4 m of the soft growth zone.

'Micro-atoll'-shaped stromatolites are also common. Micro-atolls are stromatolites with a central depression and an elevated rim, the central depressions being 10–30 cm deep and often filled with rippled oolitic sand. During maximum tidal flow, sand, shells and intraclasts swirl in the central depressions like pebbles in stream potholes. Whether the depressions are erosional features or are caused by growth of the rim has yet to be determined.

Where the channel turns southward (Fig. 1), the megaripples lose relief and the stromatolites become smaller. In this area the coalesced forms become S-shaped to form rows perpendicular to currents but with low relief, <1 m, and in plan view resemble Arabesque ornamentation. Circular micro-atolls interspersed with these shapes suggest they result from coalescence and erosion of semi-circular micro-atolls.

Staining experiments confirmed that stromatolites can grow rapidly. Three specimens were stained *in situ* with Alizarin Red S by injecting the solution into a clear inverted plastic bag cinched tightly around the base of each specimen. The bag and stain were removed after 6 h. Twenty-four h later, a layer of sand ~1 ooid thick (200 µm) had been deposited over the stained surface; sediment also adhered to vertical sides. Diver observation and underwater video documentation showed that most

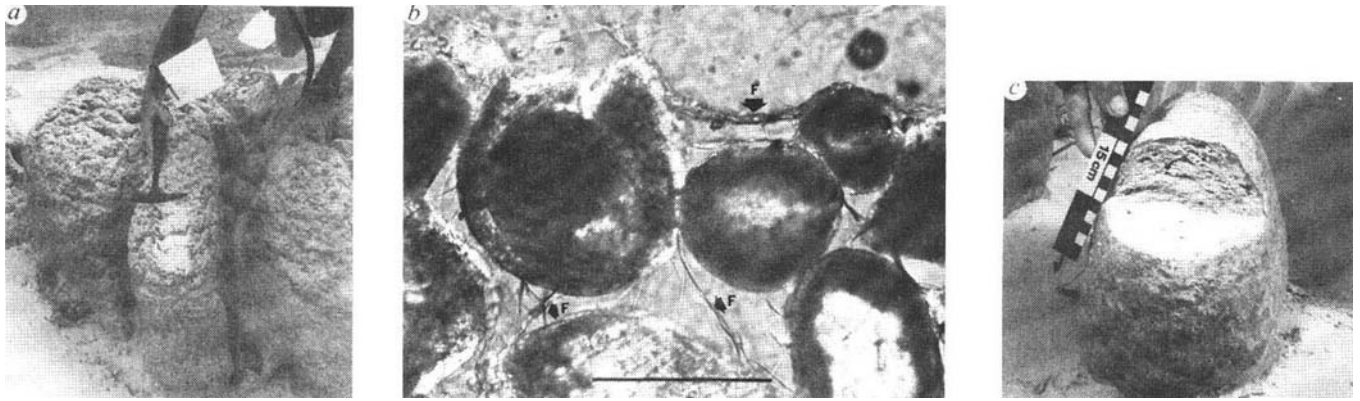


Fig. 3 *a*, Actively growing stromatolites with soft pustulate surface. *Batophora* or other epibionts are absent from surface of actively growing stromatolites. White patch below hammer caused by scraping away the soft uncemented surface layer, which is held in place by microbial filaments. Sand thrown against vertical sides by diver will stick to form a layer ~1 ooid thick. *b*, Thin-section photomicrograph of growing uncemented upper surface of pustulate stromatolite. Note tiny microbial filaments (F) responsible for trapping of grains. Large grain is an ooid. Other grains are hard peloids. Note how filaments coalesce to form horizontal mat over growth surface (arrow). Growth of the internal filaments is normal to the surface. Section prepared by preserving pustule in ethyl alcohol, followed by desiccation in oven and vacuum impregnation with polyester resin. Specimen photographed under crossed polars using a gypsum plate to emphasize microbial filaments. Portion of ooid grain was plucked during preparation but algal mat shows original surface of the ooid. Scale bar, 200 μm . *c*, Underwater photograph of smooth surface of streamlined stromatolite which has been sawed with a hand saw. Note internal laminations, lack of pustules as shown in *a*, and lack of crenulations in convex-upward laminae. Ooid sand megaripple in background partially covered specimen 24 h later.

stromatolites are dusted with ooid sand during each tidal cycle and sand thrown against vertical surfaces by a diver sticks instantly to form a layer. A shell (*Strombus gigas*) from the base of a 50-cm-high specimen at a depth of 8 m gave a ^{14}C age of 480 ± 50 yr BP. The green alga *Batophora* sp., thought by Dravis³ to be significant in stromatolite growth, was not found on the stromatolites studied here. The organic mat responsible for trapping suspended ooids and biogenic sands is a complex consortium of blue-green algae, specially adapted diatoms and as yet unidentified plants and organisms. The composition probably varies with season and time when buried columns are re-exposed. It is significant that the micro-ecosystem has adapted to high-energy conditions with life forms having 'sticky flypaper' surfaces that trap sand grains, which, in turn, on stabilization, protect the mat from further bombardment by suspended grains.

In areas near the sides of the channels, ooid sands have been stabilized by marine grasses, and sand saltation does not occur. The stromatolites in these areas are densely populated by grazing fish and epibionts such as *Acetabularia*, *Batophora*, *Sargassum*, *Siderastrea* spp. and a variety of unidentified sponges and crustose calcareous algae. Most have been invaded by boring pholad clams, and the boring sponge *Cliona* spp. has attacked many specimens; these stromatolites appear to be relict forms that are being destroyed by bioerosion.

The internal structures of dissected columns are strikingly similar to Precambrian forms and consist of convex-upward laminations composed of ooid and pelletal sand. Laminae are caused by a combination of variations in grain size, cementation and alignment of pores. Some columns have incorporated siliceous spicules. The laminations occur in two styles: nested small-scale convex-upward laminations that form 1- to 4-cm-wide columns, and larger convex-upward laminations that encompass the large portions of columnar stromatolites (Fig. 4). The smaller columns, often separated by unlaminated cemented sand, are superimposed on the larger scale laminations. Many laminations thin towards the edges of the stromatolite and in places are truncated by near vertical wall structure along the sides of the more columnar stromatolites (Fig. 4). Larger stromatolites have not been sampled, but it is thought they may be composed of both small and large columns as much as 1 m high or more. The maximum synoptic relief determined by using any one lamina in the stromatolite structure (Fig. 4) is only 16 cm, whereas the total exposed relief of the structure itself is

>45 cm; maximum synoptic relief of the small-scale column is <1–2 cm.

Voids are numerous, and whereas some are borings, others are constructional. Borings, whether excavated by pholads, serpulids, or sponges, are characterized by truncated grains and cement, and except for those made by pholads, they tend to be irregular in shape and orientation. Constructional voids are planar and parallel local laminations. Grains and cement are not truncated by constructional planar voids; these can form when fleshy epiphytes such as sponges and the alga *Batophora* sp. are entombed by subsequent deposition of laminated sand. The organism decays and a void, essentially parallel to depositional laminae, is formed and preserved by rapid cementation. Finer grained sand and silt may form geopetal fillings in both kinds of voids, which often contain uncemented ooids.

Marine cementation is rapid in this ooid-forming environment, and a portion of the Bahama Bank a few kilometres west of the Exuma Islands is where the process was first documented in the Bahamas¹¹. Cementation of the stromatolites is similar to that in Shark Bay¹². Uncemented ooid sand at the surface transitionally changes to hard limestone in the interior. The gradient from organically-bound sediment to rock extends from the upper surface to a depth of 4–5 cm and is best appreciated by sawing vertically into a stromatolite with a carpenter's hand saw; sawing begins easily but slows and becomes impossible beyond 30 cm as cementation increases.

The cement is mainly acicular aragonite, similar to that observed in other areas of marine cementation. In localized areas, aragonite crystals are tightly bundled to form what has been termed acicular fan druses¹³ or botryoidal aragonite¹⁴; the latter occurs as a partial filling in voids. Micrite-textured Mg calcite is patchy and is more common in finer grained geopetal internal sediments. Magnesian calcite also occurs as cryptocrystalline cement in 20- to 25- μm -diameter tubules³. Because of their habit of penetrating both pore space and grains, Dravis¹⁵ defined the tubules as chasmolithic/interolithic algae, probably of the genus *Ostreobium*. We are unsure of their actual affinity of the timing of their entry into the stromatolite matrix. They are not common in the uncemented growth zone and thus are not considered significant as primary sediment trappers. Additional study is needed to determine when and how they invade stromatolites and become diagenetically altered to micritic Mg calcite.

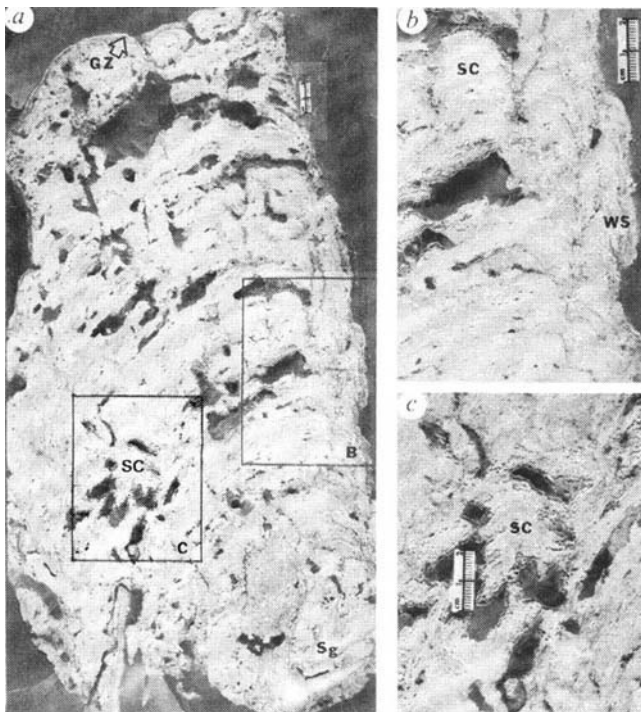


Fig. 4 Views of plastic-impregnated specimen which was cut in half to reveal internal structure. The slab is oriented so that incoming tide flowed from right to left. *a*, Slabbed plastic-impregnated specimen is 45 cm high and nucleated on a conch (*Strombus gigas*) shell (Sg). Note the small-scale convex-up laminations (SC) which form columns radiating from the base. Small-scale laminations are actually upward continuations of crenulations in larger scale laminations. Wall structure (WS) is shown in inset (*b*). Larger plastic-filled voids are borings; however, void around and below SC in inset (*c*) is thought to have been formed when a fleshy epibiont rotted away. GZ (top of *a*) shows soft uncemented pustular growth zone where grains are trapped and bound by the algal filaments shown in Fig. 3*b*. The small-scale columnar structure is thought to be caused by upward growth of pustules such as shown here. Unlaminated areas are former voids which have been filled with oolitic sand.

Bahamian subtidal stromatolites provide a new set of environmental parameters to use in interpreting spatial distribution of ancient stromatolites, orientation, water depth, salinity, constraints on growth forms and the timing of cementation and mineralization. For example, due to the Shark Bay modern analogue, most authors have regarded ancient stromatolites as intertidal in origin, although some have subtidal affinities or large size, which necessitated postulating enormous paleotidal fluctuations⁸. Modern subtidal stromatolites have been reported before^{3,6,16,17} but unfortunately seem to have had little influence on the mainstream of geological interpretation. The Bahamian examples described here are by far the deepest and largest to be reported from the Holocene and hopefully may help clarify the many debates concerning salinity, current regimen, and depositional water depth of ancient stromatolites. Ancient stromatolites, like modern ones, were not limited to intertidal or hypersaline conditions, and many authors recognized this as a fact based on geological interpretation^{8,18-20}. The Bahamian stromatolites provide an alternative for palaeoenvironmental comparative studies.

In the Holocene and probably throughout much of Phanerozoic time, grazers and borers prevented stromatolite growth and preservation in normal marine subtidal waters. During the Precambrian, however, there were no known grazing and boring organisms, and stromatolites flourished uncontested for $>2 \times 10^9$ yr. Without competition, Precambrian stromatolites freely populated enormous areas^{20,21}, and probably grew in water

depths >10 m (ref. 8). Strong tidal currents were not a necessary requirement for growth and preservation in the Precambrian before the evolution of higher life forms. We anticipate that future comparative research on these Holocene stromatolites will provide many additional clues for interpreting depositional environments and the processes that moulded the shapes and distribution of ancient stromatolites of all ages.

We are grateful for the support and facilities at Lee Stocking Island provided by the Caribbean Marine Research Center. We acknowledge the support of Mr Robert Wicklund, Ms Geraldine Wenz, and Mr John H. Perry Jr, owner of Lee Stocking Island. Partial support was provided by the Office of Undersea Research, National Oceanic and Atmospheric Administration, US Department of Commerce. Ralph Knight is thanked for piloting us. We thank Richard Wilkens, Thomas Lorensen and Tom Judy for underwater support, field and laboratory help in this unfunded research. Barbara Lidz provided editorial assistance. We also thank Sal Valestro at the Balcones Research Center of the University of Texas for dating the conch shell and R. N. Ginsburg, Richard Slater and Christopher Kendall for field assistance and stimulating discussion. Finally, we thank J. H. Hudson for guidance, equipment and chemicals for staining stromatolites *in situ* and Renato Diaz for sample and thin section preparation, and photoprocessing.

Received 11 June; accepted 22 August 1986.

1. Ritten M. G. *The Geological Aspects of the Origin of Life on Earth* (Elsevier, Brussels, 1962).
2. Berkner, L. V. & Marshall, L. C. in *The Origin and Evolution of Atmospheres and Oceans* (eds Brancaccio, P. J. & Cameron, A. G. W.) 102-126 (Wiley, New York, 1964).
3. Dravis, J. J. *Science* **219**, 385-386 (1983).
4. Logan, B. W. *J. Geol.* **69**, 517-533 (1961).
5. Logan, B. W., Rezak, R. & Ginsburg, R. N. *J. Geol.* **72**, 68-83 (1964).
6. Playford, P. E. *Western Aust. Geol. Survey A. Rep.* 1979 73-77 (1980).
7. Awramik, S. M. & Vanyo, J. P. *Science* **231**, 1279-1281 (1986).
8. Monty, C. in *Fossil Algae, Recent Results and Developments* (ed. Flügel, E.) 15-35 (Springer, Berlin, 1977).
9. Illing, L. V. *Am. Ass. Petrol. Geol. Bull.* **38**, 1-95 (1954).
10. Ball, M. M. *J. Sedim. Petrol.* **37**, 556-591 (1967).
11. Taft, W. H., Arrington, F., Haimoritz, A., MacDonald, C. & Woolheater, C. *Bull. Mar. Sci. Gulf Caribbean* **18**, 762-828 (1968).
12. Burne, R. V. & James, N. P. *12th IAS Congr.* Canberra (in the press).
13. Shinn, E. A. *Sedimentology* **12**, 109-144 (1969).
14. Ginsburg, R. N. & James, N. P. *Geology* **4**, 431-436 (1976).
15. Dravis, J. J. *J. Sedim. Petrol.* **49**, 0195-0208 (1979).
16. Ginsburg, R. N. *Proc. int. Geol. Cong. XXI, Norden 1960*, pt XXII (Copenhagen) 26-35 (1960).
17. Gebelein, C. D. *J. Sedim. Petrol.* **39**, 49-69 (1969).
18. Playford, P. E. & Cockbain, A. E. *Science* **165**, 1008-1010 (1969).
19. Hoffman, P. E. *Am. Ass. Petrol. Geol. Bull.* **58**, 856-867 (1974).
20. Cloud, P. E. *Am. J. Sci.* **240**, 363-379 (1942).
21. Walter, M. R. *Spec. Pap. Palaeontol.* No. 11 1-190 (Palaeontological Association of London, 1972).

Sex-ratio manipulation in the common opossum

Steven N. Austad & Mel E. Sunquist*

Department of Organismic and Evolutionary Biology, The Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138, USA

* Florida State Museum, University of Florida, Gainesville, Florida 32611, USA

The idea that offspring sex ratio is affected by parental capacity for reproductive investment was first developed by Trivers and Willard¹ and later elaborated by others²⁻⁶. According to this hypothesis, if the amount of parental investment⁷ differentially affects the expected reproductive success of offspring depending on offspring gender, then parents capable of exceptionally large investment should bias reproductive investment toward the sex affected most. For most polygynous mammals, male reproductive success is expected to be most enhanced by exceptionally large investment^{1,8}, but observational field studies and laboratory experiments have either supported⁹⁻¹³ or contradicted^{3,14-16} this conditional sex-allocation hypothesis. We have performed the first

experimental field study specifically to examine this hypothesis, and find that the female common opossum (*Didelphis marsupialis*), whose capacity for parental investment was artificially enhanced by dietary supplementation, produces male-biased sex ratios, in contrast to the unbiased ratios produced by control females.

The conditional sex-allocation hypothesis holds that variance in reproductive success is expected to be greater among males than among females in polygynous animals because some males have more than one mate whereas other males have no mates. Consequently, because high-quality males will have greater reproductive success than any single female, females producing high-quality offspring will have more grand offspring if they produce males. Conversely, low-quality offspring should be female, because essentially all females find mates.

The common opossum, a polygynous marsupial¹⁷, exhibits no paternal care. Young are born after 13 days' gestation in an embryonic state and complete development during 60–70 days in the maternal pouch. During a 2-year study near Calabozo in the central Venezuelan savanna, we trapped, weighed, radio-collared and released 40 adult virgin females (virgins can be identified by the condition of their pouch¹⁸) before the onset of the breeding season. Half of these females were randomly assigned to control and half to experimental groups. Two controls died before the start of breeding and are not included in the results. Experimental animals were located in their day time sleeping dens and provisioned with ~125 g sardines every 2 days from at least 2 weeks before any pouch young appeared until the end of the breeding season. Provisions were placed at the mouth of their dens near dusk when the animals become active. Provisioning was intended as a supplement to, and not a substitute for, a normal diet. Both controls and provisioned animals were recaptured, weighed and the size and sex of their pouch young recorded approximately once per month.

Pouch young can be toe-clipped and sexed visually at about 10–15 days of age when they weigh only ~1 g (<0.1% of maternal body weight). The course of pouch-young development in opossums has been extensively studied^{18–20}, so that from a combination of anatomical features the animals can be aged to within 5 days for their first month in the pouch.

We found that provisioned females have no more offspring than controls (Table 1), but their individual offspring are significantly larger at all stages of development from 20 days of age until they left the maternal pouch (Fig. 1). Thus, provisioned females exhibit greater total parental investment than controls. The recapture rate of independent juveniles supports the contention that the offspring of provisioned females survived better than controls (Table 1).

Table 1 Maternal investment in control versus provisioned females

	Control	Provisioned	P-value
Maternal weight (kg) at start of experiment ($\bar{X} \pm \text{s.e.m.}$)	1.38 $\pm$ 0.07	1.37 $\pm$ 0.07	0.71 ($t_{36} = 0.314$)
Litter size ($\bar{X} \pm \text{s.e.m.}$)	7.52 $\pm$ 0.36	7.61 $\pm$ 0.34	0.56 ($t_{36} = 0.142$, one-tail)
Pouch young mortality (%) from age <15 days to >45 days			
Males	2.3 (87)	4.5 (111)	0.30
Females	2.3 (87)	5.2 (77)	0.22
Total	2.3 (174)	4.8 (188)	(G test)
Weaned young recaptured (%)			
Males	16.4 (55)	44.4 (54)	0.001
Females	12.8 (47)	28.9 (45)	0.054
Total	14.7 (102)	37.4 (99)	0.0002 (G test)

Numbers in parentheses, total number in sample.

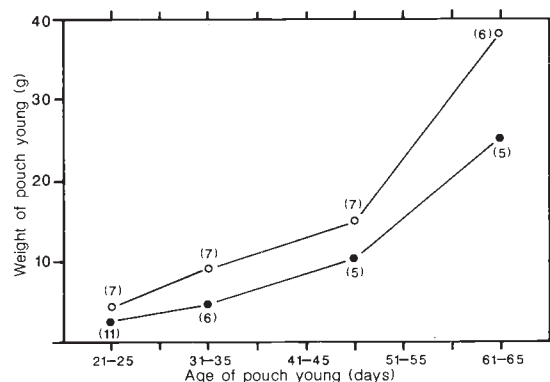


Fig. 1 Difference in weight of pouch young between provisioned (○) and control (●) females. Data from all 5-day periods during which at least 5 females from both groups were captured. Weight differences are statistically significant for all comparisons (Wilcoxon rank sum tests: 21–25 days, $P = 0.038$; 31–35 days, $P = 0.007$; 46–50 days, $P = 0.015$; 61–65 days, $P = 0.026$). No. of females in each sample in parentheses. Fifteen females from both provisioned and control groups are represented. Samples at 46–50 and 61–65 days are independent in that no female is represented in both samples.

Reproductive allocation can be sexually biased even with an unbiased sex ratio if investment is greater in individuals of one sex^{21–23}. Although there is often marked variation in size among litter mates, we found no evidence of skewed somatic investment in individuals of either sex. Mean weights of male and female litter mates are statistically indistinguishable (Wilcoxon signed-rank test for paired observations, $P > 0.2$ for controls and $P > 0.5$ for provisioned females). Neither is there a sexual bias in mortality of pouch young examined between 15 and 45 days of age (Table 1) in either provisioned or control groups (G test, $P > 0.8$ for both groups). Although maternal investment per individual is unbiased in both groups, provisioned females invest disproportionately in males by producing male-biased sex ratios. Controls, on the other hand, produce exactly as many male as female offspring (Fig. 2). The sex-ratio difference between groups is statistically significant: P (one-tail) = 0.007, binomial test. Overall, male-biased litters were produced by 55% of control females and 76% (13 of 17) of provisioned females with a sex-ratio bias.

The mechanism of this apparent sex-ratio manipulation is unknown. The sex ratios described above are those observed when pouch young were first old enough to be visually sexed. However, <2% of pouch young attached to teats died before the age at which they could be sexed, indicating that the sex ratios observed were essentially the same as those of initially attached young. Female opossums typically bear more young than can make their way to the pouch and find a functional teat¹⁸, so it is possible that males of provisioned mothers are superior to females in their capacity to migrate successfully to the pouch. Alternatively, adult females might distinguish male from female young at birth and selectively remove females as they travel toward the pouch. Another possibility is that sex-ratio manipulation occurs before birth.

For this particular sex-ratio manipulation to be an adaptive response, the amount of maternal investment would have to have a larger effect on male than female reproductive success. One reason to suspect that this may be the case is the tendency for improvement in recapture rate of independent young from provisioned versus control females to be greater for males than females ($P = 0.15$, binomial contrast; Table 1), suggesting that male survival is more affected by increased investment than is female survival. Other support for a greater effect on future male reproductive success derives from the putative relationship between adult body size and reproductive success. Among our control females, body size is uncorrelated with such fitness components as length of adult life ($r = 0.088$, $t_9 = 0.250$, $P > 0.5$),

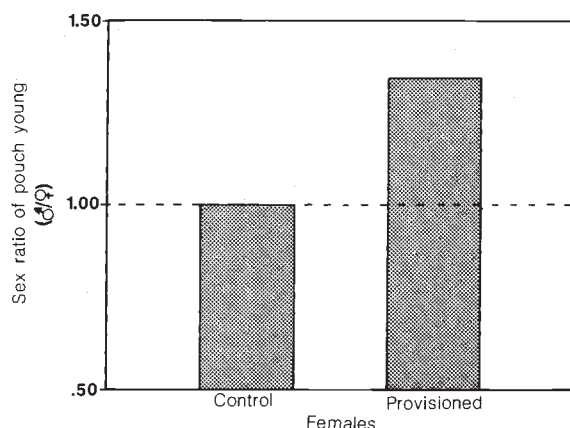


Fig. 2 Sex ratios of pouch young from control and provisioned females. Sex-ratio difference is significant ($P(\text{one-tail})=0.007$, binomial test). Control females, $n=256$; provisioned females, $n=270$.

litter size ($r=0.135$, $t_{16}=0.670$, $P>0.5$), number of litters produced ($r=-0.065$, $t_9=0.195$, $P>0.5$) or size of individual offspring ($r=-0.370$, $t_9=1.19$, $P>0.25$). On the other hand, male body size could well be related to mating success. Male common opossums have large overlapping home ranges, so that each male has potential access to multiple females, but must contend with many other males who also have potential access¹⁷. Males fight vigorously in the presence of a receptive female, and a high proportion of all males captured during the breeding season show wounds consistent with male-male combat (our unpublished observations), whereas females do not exhibit such wounds. Male body size is known to be correlated with the ability to win contests in some mammals²⁴⁻²⁶ including marsupials^{27,28}, and therefore plausibly correlates with access to mates.

Documentation of adaptive sex-ratio manipulation among mammals has remained elusive^{2,6,8,29}. Part of the difficulty in observational studies could be in identifying key variables correlated with capacity for parental investment. For instance, the offspring sex ratio in red deer varies adaptively with maternal dominance rank¹², but does not vary with other plausible variables such as maternal age, home range size or parity³⁰. Environmental variation could also obscure significant trends during any field study; for example, in a similar study of the Virginia opossum, food supplementation resulted in an overproduction of males during a year of normal rainfall, but failed to do so in a drought year (M.E.S., in preparation), possibly because even provisional females were not in exceptionally good condition during the drought. However it seems increasingly clear that taxonomically diverse mammal species can adjust their sex ratios as the theory predicts.

We thank Tomás Blohm for allowing us to work on his ranch and assisting us with logistical problems. We also thank Dennis Daneke and his assistants Brent Mitchell and David Manry for performing most of the fieldwork. We thank E. L. Charnov, T. H. Clutton-Brock, G. D. Dodson, J. F. Eisenberg, L. D. Marshall, J. Maynard Smith, W. R. Rice, M. J. A. Simpson and R. Thornhill for helpful comments on the manuscript. Supported by NSF grant BSR-8315125.

Received 28 March; accepted 21 August 1986.

- Trivers, R. L. & Willard, D. E. *Science* **179**, 90-92 (1973).
- Williams, G. C. *Proc. R. Soc. B* **205**, 567-580 (1979).
- Altmann, J. *Baboon Mothers and Infants* (Harvard University Press, 1980).
- Fiala, K. L. *Evolution* **35**, 898-911 (1981).
- Okansen, L. *Am. Nat.* **117**, 109-111 (1981).
- Charnov, E. L. *The Theory of Sex Allocation* (Princeton University Press, 1982).
- Trivers, R. L. in *Sexual Selection and the Descent of Man 1871-1971* (ed. Campbell, B.) 136-179 (Aldine, Chicago, 1972).
- Clutton-Brock, T. H. & Albon, S. D. in *Current Problems in Sociobiology* (ed. King's College Sociobiology Group) 223-247 (Cambridge University Press, 1982).

- Rivers, J. P. W. & Crawford, M. A. *Nature* **252**, 297-298 (1974).
- McClure, P. A. *Science* **211**, 1058-1060 (1981).
- Gosling, L. M. & Petrie, M. in *Physiological Ecology* (eds Calow, P. & Townsend, C. R.) 315-345 (Blackwell Scientific, Oxford, 1982).
- Clutton-Brock, T. H., Albon, S. D. & Guinness, F. E. *Nature* **308**, 358-360 (1984).
- Rutberg, A. T. *Am. Nat.* **127**, 89-94 (1986).
- Verme, L. J. *J. Wildl. Mgmt.* **47**, 573-582 (1983).
- Simpson, M. J. & Simpson, A. E. *Nature* **300**, 440-441 (1982).
- Silk, J. B. *Am. Nat.* **121**, 56-66 (1983).
- Sunquist, M. E., Austad, S. N. & Sunquist, F. J. *Mammal* (in the press).
- Petrides, G. A. J. *Mammal* **30**, 364-378 (1949).
- Reynolds, H. C. *Univ. Calif. Publ. Zool.* **52**, 223-284 (1952).
- Gardner, A. L. in *Wild Mammals of North America* (eds Chapman, J. A. & Feldhamer, G. A.) 3-36 (Johns Hopkins University Press, 1982).
- Reiter, J., Stinson, N. L. & LeBoeuf, B. J. *Behav. Ecol. Sociobiol.* **3**, 337-367 (1978).
- Dittus, W. P. J. *Behaviour* **69**, 265-301 (1979).
- Maynard Smith, J. *Behav. Ecol. Sociobiol.* **7**, 247-257 (1980).
- Geist, V. *Mountain Sheep* (University of Chicago Press, 1971).
- Clutton-Brock, T. H., Guinness, F. E. & Albon, S. D. *Red Deer: Behaviour and Ecology of Two Sexes* (Cambridge University Press, 1982).
- Fleming, T. H. J. *Mammal* **55**, 543-561 (1974).
- Kaufmann, J. H. *Am. Zool.* **14**, 51-62 (1974).
- Johnson, C. N. *Austral. Wildl. Res.* **10**, 25-31 (1983).
- Myers, J. H. *Am. Nat.* **112**, 381-388 (1978).
- Clutton-Brock, T. H., Albon, S. D. & Guinness, F. E. *Nature* **289**, 487-489 (1981).

Wormy mice in a hybrid zone

Richard D. Sage*†, Donald Heyneman‡, Kee-Chong Lim‡ & Allan C. Wilson*

* Department of Biochemistry and † Museum of Vertebrate Zoology, University of California, Berkeley, California 94720, USA

‡ Department of Epidemiology and International Health, University of California, San Francisco, California 94143, USA

As one approach to analysing the genetic barriers between species, we studied the numbers and types of parasitic worms in two species of house mice (*Mus musculus* and *M. domesticus*) and in their natural hybrids. Where the ranges of these two species meet in southern Germany, there is a zone of hybridization less than 20 kilometres across¹, in which about 98% of the mice have backcross genotypes. Fourteen of the 46 mice tested from within the zone have over 500 pinworms per gut, a number far exceeding the mean of 40 per gut for other mice inside and outside the zone. Other nematodes have a similar, non-random distribution. The number of mice bearing 9 or more tapeworms per gut is also excessive in the hybrid zone. These extraordinarily wormy mice may be unusually susceptible to parasitism; the different species may have different genes for resistance, and recombinant backcross animals may lose both². Our findings support the view that the hybrid populations may have reduced fitness and thereby act as a genetic sink, interfering with the flow of genes between the two species^{1,3}. The possibility that environmental or ecological peculiarities in the zone of hybridization make the mice more liable to infection is not supported.

Ninety-three mice, caught in barns and houses at 22 localities in southern Germany and adjacent Austria, were necropsied for worms in their tissues and intestinal tracts. These mice are part of a larger collection of about 300 individuals from 30 localities along a 200-km transect extending between Braunau and Augsburg. This transect encompasses the reported zone of hybridization between the eastern *M. musculus* and western *M. domesticus*^{4,5}. A partial genetic characterization of mice from these 30 localities has appeared¹. Animals were trapped alive during October 1984 (42 mice) and June 1985 (51 mice) and killed within 2 days of capture. All but 5 of the 93 mice were adults, weighing more than 12 g, and 40 were males. The carcasses were preserved in 10% neutral-buffered formalin for 1-2 months, soaked in water, and stored in 70% ethanol. The entire gastrointestinal tract was slit and examined, with associated organs, under a dissecting microscope.

The genotype of each mouse at four enzyme loci was determined by starch gel electrophoresis¹. These two species have

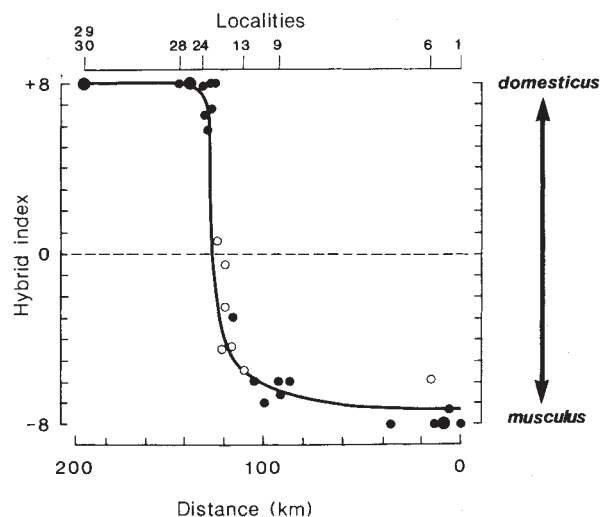


Fig. 1 Mean hybrid indices of 30 mouse populations based on electrophoresis of four diagnostic proteins. The samples come from localities 1–30 lying along a line extending between Braunau, Austria (1), and Augsburg, Germany (29, 30; ref. 1). Locality numbers of some of the 22 places from which mice were autopsied for parasites are shown in Fig. 1, there is a nearly complete replacement of one species by the other across a distance of less than 20 km. The mitochondrial DNA, chromosomes, osteology, pelage coloration and hybrid sterility alleles also show this abrupt transition (R.D.S., unpublished).

fixed allelic differences at loci for esterases 1 and 10, mannose phosphate isomerase-1 and nucleoside phosphorylase-1. A hybrid index (HI) based on the genotypes at these four loci was computed for 200 mice, including the 93 animals autopsied for parasites. As shown in Fig. 1, there is a nearly complete replacement of one species by the other across a distance of less than 20 km. The mitochondrial DNA, chromosomes, osteology, pelage coloration and hybrid sterility alleles also show this abrupt transition (R.D.S., unpublished).

We found and identified more than 10,840 worms in the 93 mice. From Table 1, which gives the average number of worms per mouse at the 22 localities, it is evident that the highest worm burdens occur in the hybrid zone. Here, between localities 13 and 23, the average number of nematodes per mouse is greater than 197, the standard error being 32. By contrast, in the east the mean is 59 (± 23) and in the west it is only 15 (± 4). The caecal pinworms (*Aspiculuris* and *Syphacia*), which comprised 90% of the nematodes (Table 1), and other genera (*Protospirura*, *Trichocephalus* and *Trichuris*) also occurred in highest numbers in hybrid mice. A similar trend is apparent for cestode infections

Table 1 Numbers of worms per gut in 93 wild mice

Locality*	Hybrid index	No. mice	Mean abundance (maximum)	
			Nematodes†	Cestodes‡
Eastern region (<i>M. musculus</i>)				
1	-8	3	4.3 (10)	1.7 (3)
2	-7.3	2	7.5 (15)	0
4	-8	2	16 (30)	0
6	-6.0	8	100+ (500+)	1.4 (3)
9	-6.6	4	80 (151)	0.8 (1)
12	-6.0	2	25 (30)	0
Hybrid zone				
13	-5.5	2	255+ (500+)	0
14	-3.0	2	20 (40)	0
15	-4.3	5	423+ (520+)	11 (22)
16	-0.5	4	170+ (500+)	0.6 (2)
17	-4.5	7	79+ (500+)	0
18	-2.5	3	350+ (500+)	3.0 (9)
19	+0.6	13	243+ (620+)	5.8 (30)
20	+6.7	4	44 (100)	1.0 (4)
23	+5.7	6	126 (200)	0
Western region (<i>M. domesticus</i>)				
24	+6.5	7	2.0 (6)	0
25	+7.8	4	22 (30)	0
26	+8	3	8.3 (14)	0
27	+8	1	8	11
28	+8	2	28 (52)	0.5 (1)
29	+8	6	8.3 (15)	0.5 (2)
30	+8	3	54 (100)	0

* The locality numbers and the mean hybrid indices are given in ref. 1, Table 1.

† The five genera represented and total numbers of specimens found are: *Aspiculuris* (>5,740), *Syphacia* (>3,865), *Protospirura* (485), *Trichocephalus* (512) and *Trichuris* (65).

‡ The principal cestode genera (and the total numbers found) were *Hymenolepis* (97), *Catenotaenia* (64) and *Taenia* (14).

(Table 1): mice in the hybrid zone have an average of 3.2 (± 0.97) worms, whereas the eastern and western mice have 0.9 (± 0.23) and 0.6 (± 0.13) respectively.

The number of worms per mouse varies widely, from none to >620 (Table 1, Fig. 2). Table 2 shows that the distributions of worm burdens per mouse are highly skewed for each of the three regions, the variance exceeding the mean in nearly every case. Such skewed distributions are characteristic for parasites within their host populations^{6,7}. The ratio of the variance to the mean is highest for the hybrid zone. This observation, in conjunction with the evidence shown in Fig. 3 and discussed below, may imply an association between worm numbers and particular genotypes.

We designate the 15 mice bearing >500 nematodes as 'wormy' and the remaining 78 mice (<250 nematodes) as 'normal' (most of the normal mice carry <50). The pattern of nematode burdens is associated with particular protein genotypes (Fig. 3). Fourteen of the 15 wormy mice have hybrid genotypes. Only one of these heavily parasitized mice has the genotype of a pure *M. musculus*. A χ^2 analysis of pure versus hybrid genotypes and wormy (>500 nematodes) versus normal (<250) mice gives a highly significant result ($\chi^2 = 9.6$, d.f. = 1, $P < 0.005$). Thus there are significantly fewer pure mice, and more hybrids than expected, that are heavily parasitized with nematodes.

Cestode infections show a similar pattern (Fig. 2b), although the maximum number of parasites per mouse is not as great as for nematodes. There are 84 normal animals with from zero to four cestodes, and nine wormy mice with heavy infections ranging from 9 to 30 worms. Eight of these wormy mice have hybrid genotypes (HI = -4 to +4), while one pure *domesticus* mouse had 11 cestodes. A χ^2 test of pure versus hybrid genotypes and

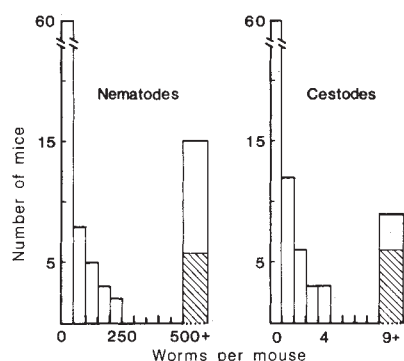


Fig. 2 The distribution of parasite burdens (number of worms per mouse) among 93 wild house mice captured at the 22 localities cited in Fig. 1: a, nematodes b, cestodes. Hatched columns, six mice that had excessive numbers of both nematodes (>500) and cestodes (>9).

Table 2 Statistical analysis of worm numbers per mouse

Type of worm	Variance divided by the mean		
	Eastern region (n = 21)	Hybrid zone (n = 46)	Western region (n = 26)
Nematodes			
<i>Syphacia</i>	263	368	50
<i>Aspicularis</i>	51	343	30
<i>Protospirura</i>	3.9	84	1.0
Trichurids	9.1	160	18
Cestodes			
<i>Hymenolepis</i>	6.9	25	10
Other taenioids	1.5	12	0.8

wormy (>8 cestodes) versus normal (<5) mice gives a significant result ($\chi^2 = 4.1$, d.f. = 1, $P < 0.05$). We infer that hybrid mice can be more heavily parasitized than genetically pure *M. musculus* or *M. domesticus*.

Excessive worminess seems not to be related to the sex of the mouse, as reported elsewhere⁸. Seven of the 15 mice with >500 nematodes, and three of the six animals most heavily infected with cestodes, were males. Also, the heavy infections are widely distributed within the hybrid zone and present during both collecting periods, suggesting that the causative agent does not come from some unique ecological or temporal factors in a particular mouse colony. No differences in habitat or mouse abundance were obvious between the collecting sites within the hybrid zone and in the adjacent regions. The 15 mice with the highest nematode burdens come from seven cities (Fig. 1), and the nine animals with heavy cestode infestations were collected at four locations (numbers 15, 18, 19 and 27).

Neither the high mean levels of nematodes in the hybrid zone ($\bar{x} = 198$) nor the extraordinarily high maximal counts (>500) found in some mice have been found in other sampled populations of wild *M. musculus*⁹ or *M. domesticus*¹⁰⁻¹⁴. Although the pathological effects of these heavy worm burdens on wild mice have not been studied, laboratory experiments implicate pinworms in causing rectal prolapse, enteritis, liver granulomas and poor weight gains¹⁵, as well as in altering immune responses to other disease agents¹⁵. Although the worm burdens are clearly not fatal to these wild mice, they seem likely to be deleterious.

The diversity of life cycles among these parasites is notable. The caecal pinworms and some of the intestinal *Hymenolepis* are transmitted directly from mouse to mouse but others have complex life cycles involving other animals as intermediate hosts. The epidemiological conditions needed to account for the worminess of these hybrids must therefore be diverse. The implication is that some of the mice are peculiarly susceptible to worm infections.

Because nearly all the wormy mice are hybrids, we suggest that the increased parasitism has a genetic basis. Reviews of the literature on resistance by small mammals to parasites¹⁶⁻²⁰ give many examples of genetically-based resistance to both uni- and multicellular parasites, including some of the helminths we found in the wild mice (*Aspicularis*, *Hymenolepis*, *Trichuris*, and *Taenia*). The molecular and cellular mechanisms that mediate resistance to worms are diverse and are known to include immune-response and histocompatibility genes in mice^{16,19,20}. It has been shown that the major genes for susceptibility to the bacterium *Salmonella typhimurium* reside on different chromosomes in *M. domesticus* and *M. musculus*², which implies different genetic mechanisms of resistance. Resistance to *Trichuris muris* and *Trichinella spiralis* behave as dominant characters. The reviews¹⁶⁻²⁰ emphasize the ubiquity and genetic simplicity of natural resistance.

These known cases of mouse resistance to parasites suggest that simple genetic models could explain our field observations. One possible model is that helminth resistance resides in

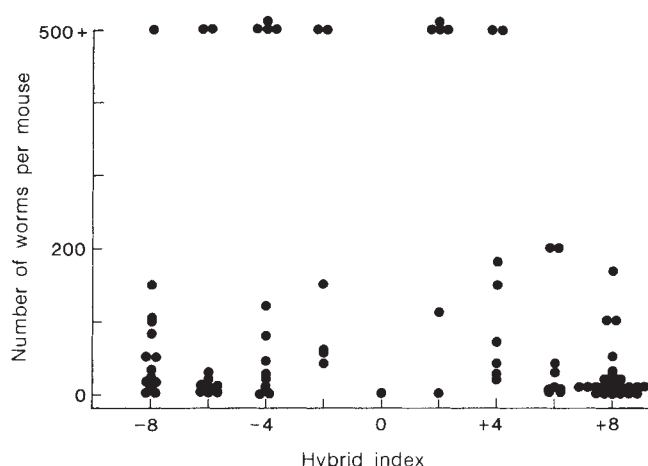


Fig. 3 Nematode burdens (number of worms per mouse) in relation to mouse genotype. The hybrid index is based on electrophoresis of four diagnostic proteins. Pure *M. musculus* has an index of -8; pure *M. domesticus* has an index of +8.

different genes in *M. musculus* and *M. domesticus*, and that wormy mice have recombinant genotypes. If the resistance genes are dominant in the wild, as in the laboratory, then the susceptible, wormy mice would arise after the initial hybridization events, from a mating between recombinant animals. Alternatively, resistance could be conferred by a combination of alleles at two linked loci and the alleles conferring resistance could differ in the two species. Recombination would break up advantageous combinations of alleles. Parental and F_1 genotypes are rare in the hybrid zone; nearly all the mice have backcross genotypes¹. Thus, these hybrid populations have genetic constitutions compatible with either of these genetic models.

The initial field observation (R.D.S.) which led to this quantitative analysis of helminth parasitism was the apparently greater numbers of fleas on the recently killed, hybrid mice, in comparison to non-hybrids. Conversations with colleagues who study hybridization in other species indicate that hybrids may frequently have more parasites than the adjacent parental populations (for example, gall aphids on cottonwoods, gill parasites in stone crabs, and intestinal worms in two genera of salamanders; T. Bert, T. Kocher, S. Sessions and T. Whitham, personal communication). If future studies verify higher-than-normal levels of parasitism and 'disease' in other hybrid zones, they would encourage the view that the loci which change fastest in response to selection after two populations separate are often those controlling resistance to parasites²¹.

We thank I. Butler, K. Damme, J. Klein, R. Kraft, F. Seidl, H. Tichy, F. Pirchner and the people who allowed us to trap mice in their buildings, P. Bateson, W. Hamilton, F. James, A. Kuris, E. Mayr, E. Prager, J. Shields and S. Sweet for discussion about the manuscript, and the NSF for financial support.

Received 15 April; accepted 8 September 1986.

1. Sage, R. D., Whitney, J. B. III & Wilson, A. C. *Curr. Top. Microbiol. Immun.* **127**, 75-85 (1986).
2. O'Brien, A. D., Weinstein, D. L., D'Hoostelaere, L. D. & Potter, M. *Curr. Top. Microbiol. Immun.* **127**, 309-318 (1986).
3. Barton, N. H. *Heredity* **44**, 277-278 (1980).
4. Kraft, R. *Saugetk. Mitt.* **32**, 1-12 (1985).
5. Zimmermann, K. *Zool. Jb. Abt. Syst. (Oekol. Geogr. Tier.* **78**, 301-322 (1949).
6. Crofton, H. D. *Parasitology* **62**, 179-194 (1971).
7. Anderson, R. M. & Gordon, D. M. *Parasitology* **85**, 373-398 (1982).
8. Sage, R. D. in *The Mouse in Biomedical Research* Vol. 1 (eds Foster, H. L., Small, J. D. & Fox, J. G.) 39-90 (Academic, New York, 1981).
9. Pleshchev, V. S. & Kozlov, A. N. *Mater. Nauchn. Konf. Vses. O. Gel'mint.* **30**, 135-139 (1978).
10. Elton, C. *Parasitology* **26**, 107-111 (1934).
11. Whitaker, J. O. *Proc. Indiana Acad. Sci.* **79**, 441-448 (1970).
12. Clark, D. E. *Proc. Indiana Acad. Sci.* **80**, 495-500 (1970).
13. Behnke, J. M. *J. Helminth.* **49**, 85-90 (1975).
14. Singleton, G. R. *Aust. Wildl. Res.* **12**, 437-445 (1985).
15. Wescott, R. B. in *The Mouse in Biomedical Research* Vol. 2 (eds Foster, H. L., Small, J. D. & Fox, J. G.) 373-384 (Academic, New York, 1982).

16. Wakelin, D. *Adv. Parasit.* **16**, 219–308 (1978).
17. Skamene, E., Kongshaun, P. A. & Landy, M. (eds) *Genetic Control of Natural Resistance to Infection and Malignancy* (Academic, New York, 1980).
18. Albright, J. F. & Albright, J. W. *Contemp. Top. Immunobiol.* **12**, 1–52 (1984).
19. Wassom, D. L., David, C. S. & Gleich, G. J. in *Genetic Control of Natural Resistance to Infection and Malignancy* (eds Skamene, E., Kongshaun, P. A. L. & Landy, M.) 75–82 (Academic, New York, 1980).
20. Wakelin, D. in *Ecology and Genetics of Host-Parasite Interactions* (eds Rollinson, D. & Anderson, R. M.) 39–54 (Academic, London, 1985).
21. Hamilton, W. D. in *Population Biology of Infectious Diseases* (eds Anderson, R. M. & May, R. M.) 269–296 (Springer, Berlin, 1982).

A single micromanipulated stem cell gives rise to multiple T-cell receptor gene rearrangements in the thymus *in vitro*

Gwyn T. Williams*, Rosetta Kingston*, Mike J. Owen†, Eric J. Jenkinson* & John J. T. Owen*

* Department of Anatomy, Medical School, Vincent Drive, Edgbaston, Birmingham B15 2TJ, UK

† ICRF Tumour Immunology Unit, Zoology Department, Gower Street, London WC1E 6BT, UK

The extensive range of specificities of T-cell receptors is generated, as for immunoglobulins, by rearrangement of genetic information¹. Much valuable information about rearrangement processes has been inferred by comparing DNA from (monoclonal) lymphoid lines with germ-line DNA^{2,3} and, for B cells, from rearrangements in some Abelson murine leukaemia virus-transformed cell lines^{4,5}. However, because it is difficult to isolate and grow precursor populations, it has not proved possible to study rearrangements occurring in normal untransformed cells *in vitro*. Here we show that a single T-cell precursor colonizing an alymphoid thymus lobe in organ culture can generate multiple receptor β -chain gene rearrangements. These observations provide unequivocal evidence for the intra-thymic diversification of the T-cell repertoire. They also offer the possibility of investigating rearrangement and its control in the clonal progeny of a single normal T-cell precursor without the perturbations involved in the use of viral transformation or the production of T-cell hybridomas.

T cells require both specific antigen and self major histocompatibility (MHC) determinants^{6,7} for activation. They possess a cell surface receptor made up of α - and β -chains^{8,9} encoded by genes which are rearranged and expressed only in T lymphocytes^{10,11}. The rearrangement of receptor β -chain genes during mouse fetal development has been monitored by hybridizing β -chain gene probes to restriction enzyme-digested DNA from cells at different stages of development^{12,13}. The proportion of thymic lymphocytes showing rearrangement of the β -chain genes is very small at day 14, then increases progressively. Several interpretations of these observations are possible: (1) gene rearrangement may occur in the thymus; (2) rearranged populations may proliferate preferentially in the thymus; (3) there may be progressive recruitment of rearranged cells into the thymus; or some combination of these processes. To distinguish between these possibilities we used a recently developed technique¹⁴ to recolonize an alymphoid thymic lobe with a single blast cell derived from 14-day fetal thymus. Successfully recolonized lobes yield up to 1.1×10^5 lymphocytes, and contain various T-cell marker defined sub-populations, all derived from the single donor cell¹⁴. If rearrangement had already occurred at the 14-day stage, either within the thymus or at a pre-thymic level, one or two rearranged bands would be detected. If no rearrangement had occurred by the 14-day stage or during organ culture a strong germ-line pattern would still be evident, while

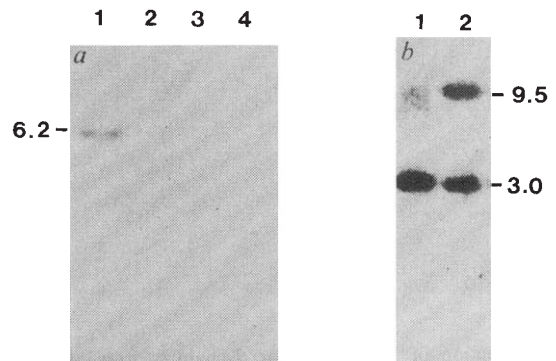


Fig. 1 Southern blot analyses of T-cell receptor β -chain genes in the populations produced in thymus lobes recolonized by a single T-cell precursor. *a*, *Pvu*II digestion of: lane 1, liver DNA; lanes 2–4, progeny of single cells; *b*, *Hind*III digestion of: lane 1, progeny of single cell; lane 2, liver DNA. A lymphoid CBA (H-2^K) thymus lobes were prepared and each was recolonized by a single T-cell precursor taken from a 14 day of gestation BALB/c thymus lobe as described previously¹⁴. *a*, DNA was prepared from the lymphoid cells produced from single precursors, digested with *Pvu*II, separated on an agarose gel, Southern blotted and hybridized with a ³²P-labelled probe for the constant region of the β -chain gene. Filters were washed at high stringency before autoradiography. Liver DNA (β -chain gene complex in germ-line state) was analysed in parallel. The two constant region genes are not resolved and migrate with an apparent size of 6.2 kb (liver DNA, lane 1). The substantial reduction in these bands in the progeny of single cells, with no dominant rearranged band, indicates that multiple rearrangements have occurred (lanes 2–4). Five other lobes analysed all showed similar results (data not shown). The filter was rehybridized with a probe for β_2 -microglobulin, confirming that about the same amount of DNA was present in all tracks. *b*, DNA from the progeny of a single cell (lane 1) and liver DNA (lane 2) were digested with *Hind*III and analysed as above. In the germ-line state (lane 2) the *C β 1* restriction fragment migrates with an apparent size of 9.5 kb, and the *C β 2* fragment with an apparent size of 3.0 kb. In the DNA from the single cell progeny the germ-line *C β 1* fragment has been sharply reduced and a smear of bands has appeared. This indicates that heterogeneous multiple rearrangements have occurred within the thymus.

Methods. A lymphoid lobes were obtained after culture in the presence of deoxyguanosine (dGuo) and associated with single micromanipulated precursors for 24 h in hanging drops formed in Terasaki plates, with subsequent transfer to organ culture for 11–13 days for optimal expansion of successfully recolonized lobes. This produces lobes exclusively recolonized by progeny of the single donor precursor¹⁴. At the end of the organ culture period lymphoid cells were released from individual thymus lobes by gentle teasing with fine cataract knives and the resultant suspension was washed in Tris-buffered Earle's saline and frozen at -80°C before preparation of DNA. For *Pvu*II digestion DNA was prepared by proteinase K digestion in 0.5% SDS followed by phenol-chloroform extraction and extensive dialysis. After *Pvu*II digestion to completion, the DNA was precipitated with ethanol and separated by electrophoresis on a 0.8% agarose gel. For *Hind*III digestion the frozen cell pellet was digested with proteinase K in 0.5% sodium lauryl sarkosinate and 1% low gelation temperature agarose (Bio-Rad). The agarose was allowed to solidify at 0°C and detergent was removed by incubation with several changes of 10 mM Tris 1 mM EDTA pH 8.0 (G.T.W., manuscript in preparation). The DNA was digested with *Hind*III in the presence of 0.1% nuclease-free bovine serum albumin and separated on a 0.6% agarose gel (normal gelation temperature). Gels were blotted onto Zeta-probe (Bio-Rad) membranes and hybridized with an oligo-labelled¹⁹ *C β* probe derived from plasmid 86T1¹¹. Filters were washed at high stringency and autoradiographed for 10 days using Kodak XAR-5 film and two intensifying screens (Dupont Cronex Lightning-plus).

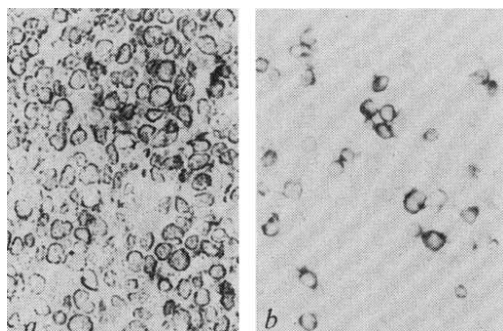


Fig. 2 Detection of T-cell receptor β -chain expression in the progeny of a single T-cell precursor. *a*, Labelling with anti-Lyt 2.2; *b*, labelling with F-23.1, a monoclonal antibody that recognizes the products of a V_{β} gene family present on approximately 25% of peripheral T cells¹⁵. A lymphoid CBA (H-2^k, Lyt 2.1) thymus lobes were recolonized by a single BALB/c (H-2^d, Lyt 2.2) precursor (Fig. 1 legend). After 11–13 days of organ culture individual lobes were snap frozen and four frozen sections prepared for indirect immunoperoxidase labelling. Two sections taken from the same lobe are shown. *a*, Labelling with monoclonal anti-Lyt 2.2 (NEN); most lymphoid cells are labelled, indicating that they have undergone differentiation in terms of T-cell marker expression and that they are derived from the single BALB/c (Lyt 2.2) precursor; as shown in our previous work, dGuo-treated lobes recolonized by a single cell contain T cells exclusively of donor type^{13,20}. *b*, The labelling with F-23.1 demonstrates productive rearrangement in a proportion of cells, but does not provide a quantitative estimate of the extent of T-cell receptor β -chain diversity, as the proportion of receptor-positive cells is unknown. However, the observed proportion of F-23.1⁺ cells is fully consistent with the heterogeneous multiple β -chain gene rearrangements shown in Fig. 1. The large reduction in the C_{β} germ-line bands shown in Fig. 1 suggests that most of the F-23.1⁺ cells have also rearranged the β -chain gene locus, although not all of these rearrangements may be productive and a proportion of cells may not synthesize any β -chain. The F-23.1 labelling might detect surface and/or cytoplasmic β -chain expression²¹ (in a recent study using a double immunofluorescence procedure we have found that, in newborn thymus, approximately equal proportions of cytoplasmic F-23.1⁺ and surface F-23.1⁺ cells are present²²). As further proof that the progeny of a single cell can proceed to productive receptor expression, lymphoid CBA lobes (Thy 1.2) were recolonized by single AKR (Thy 1.1) precursors. Double immunofluorescence labelling with anti-Thy-1.1 and F-23.1 showed that 90% of cells were Thy-1.1⁺ and 3.3% of these were surface F-23.1⁺, a value similar to that observed in non-manipulated thymus organ cultures¹³.

multiple rearrangements in the progeny of the single cell in the organ-cultured thymus would reduce the germ-line bands without producing any dominant rearranged bands.

The lymphocyte DNA from several lobes, each recolonized by a single cell, was prepared and analysed separately by restriction digestion, gel electrophoresis, Southern blotting and hybridization with radiolabelled β -chain constant region probes (Fig. 1). For *Pvu*II-digested DNA the intensity of the 6.2-kilobase (kb) band corresponding to the germ-line ($C_{\beta 1}$ and $C_{\beta 2}$) genes decreased markedly in lymphocytes from recolonized lobes, indicating extensive multiple rearrangements (Fig. 1*a*). In *Hind*III-digested liver DNA a 9.5-kb band corresponds to $C_{\beta 1}$ and a 3.0-kb band to $C_{\beta 2}$. Rearrangement to either β -chain constant region reduces the intensity of the 9.5-kb band through deletion or rearrangement of $C_{\beta 1}$, but does not affect the 3.0-kb fragment, which thus acts as an internal marker^{11,12}. The 9.5-kb ($C_{\beta 1}$) band is sharply reduced in the DNA of the progeny of the single precursor showing that the cells have

undergone multiple β -chain gene rearrangements (Fig. 1*b*). The heterogeneity of the rearrangements in the lymphoid cells is confirmed by the appearance of a smear of rearranged bands. All nine lobes analysed showed multiple rearrangements.

To demonstrate productive β -chain gene rearrangements from a single thymic stem cell we used a monoclonal antibody (F-23.1), which reacts with a sub-population of β -chain variable regions¹⁵ to look for expression of β -chain. A significant proportion of the cells produced from single precursors express T-cell receptor β -chains (Fig. 2), so many of the gene rearrangements detected are indeed productive.

The recolonization of alymphoid thymic lobes by single micromanipulated precursors has allowed us to provide a definitive demonstration of gene rearrangement involved in the generation of T-cell receptor β -chain diversity in the thymus. Neither progressive recruitment nor preferential proliferation of cells with rearranged β -chain genes could account for these observations as neither is possible in this system. This is the first *in vitro* demonstration of generation of diversity in any receptor molecule of the immune system from a single untransformed cell. Since these results show that the thymic environment is sufficient to induce β -chain gene rearrangements in T-cell precursors, they argue against the need for pre-thymic rearrangement as suggested by some functional studies^{16,17}. Furthermore, 14-day thymus cells, although capable of proliferation *in vitro*, do not undergo rearrangement outside the thymus¹⁸. Thus the thymic environment is important in initiating the rearrangement process. This system provides a method for investigating intercellular interactions in the thymus, as well as the molecular mechanisms which produce gene rearrangements during T-cell differentiation. The technique should also prove invaluable in identifying pre-thymic T-cell precursors in normal and transformed populations by their capacity to undergo receptor gene rearrangement when introduced into a thymus. The pattern of rearrangement and β -chain utilization obtained when progressively more mature cells are re-introduced into the thymus may also be analysed using this approach.

This work was supported by the MRC (E.J.J., R.K. and J.J.T.O.), the Rowbotham Bequest (G.T.W., E.J.J.) and Wellcome Trust (G.T.W.). We thank Dr M. M. Davies for clone 86T1, Dr M. Bevan for monoclonal antibody F23.1, Dr K. Docherty, Mr I. Phillips and Dr F. H. C. Franklin for help and advice, Alison Orchard for assistance with the photography and Claire Hundley for typing the manuscript.

Received 24 April; accepted 28 August 1986.

- Hood, L., Kronenberg, M. & Hunkapiller, T. *Cell* **40**, 225–229 (1985).
- Tonegawa, S. *Nature* **302**, 575–581 (1983).
- Kronenberg, M., Siu, G., Hood, L. & Shastri, N. A. *Rev. Immun.* **4**, 529–591 (1986).
- Lewis, S., Rosenberg, N., Alt, F. & Baltimore, D. *Cell* **30**, 807–816 (1982).
- Sugiyama, H. *et al. Nature* **303**, 812–815 (1983).
- Zinkernagel, R. M. & Doherty, P. C. *Nature* **248**, 701–702 (1974).
- Seidman, J. G. & Leder, P. *Nature* **276**, 790–795 (1978).
- Allison, J., McIntyre, B. & Block, D. J. *Immun.* **129**, 2293–2300 (1982).
- Haskins, K. *et al. J. exp. Med.* **157**, 1149–1169 (1983).
- Yanagi, Y. *et al. Nature* **308**, 145–149 (1984).
- Hedrick, S. M., Cohen, D. I., Nielsen, E. A. & Davis, M. M. *Nature* **308**, 149–153 (1984).
- Snodgrass, H. R., Kisielow, P., Kiefer, M., Steinmetz, M. & von Boehmer, H. *Nature* **313**, 592–595 (1985).
- Owen, M. J., Jenkinson, E. J., Williams, G. T., Kingston, R. & Owen, J. J. T. *Eur. J. Immun.* **16**, 875–878 (1986).
- Kingston, R., Jenkinson, E. J. & Owen, J. J. T. *Nature* **317**, 811–813 (1985).
- Staerz, U. D., Rammensee, H. G., Benedetto, J. & Bevan, M. J. *J. Immun.* **134**, 3994–4000 (1985).
- Morisey, P. J., Kruisbeck, A. M., Shanan, S. O. & Singh, H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2003–2007 (1982).
- Chervenak, R., Cohen, J. J. & Miller, S. D. *J. Immun.* **131**, 1688–1692 (1983).
- Palacios, R. & von Boehmer, H. *Eur. J. Immun.* **16**, 12–19 (1986).
- Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6–13 (1983).
- Jenkinson, E. J., Franchi, L. L., Kingston, R. & Owen, J. J. T. *Eur. J. Immun.* **12**, 583–587 (1982).
- Farr, A. G., Anderson, S. K., Marrack, P. & Kappler, J. *Cell* **43**, 543–550.
- Owen, J. J. T., Kingston, R. & Jenkinson, E. J. *Immunology* **59**, 23–27 (1986).

Extensive proliferation of mature connective-tissue type mast cells *in vitro*

Tatsutoshi Nakahata*, Toshimi Kobayashi*, Akira Ishiguro*, Kohichiro Tsuji*, Kuniaki Naganuma*, Osaaki Ando*, Yoshio Yagi*, Kenji Tadokoro† & Taro Akabane*

* Department of Pediatrics, Shinshu University School of Medicine, Asahi 3-1-1, Matsumoto 390, Japan

† Department of Medicine and Physical Therapy, University of Tokyo, Tokyo 113, Japan

There are two phenotypically distinct subpopulations of mast cells in rodents: connective tissue-type mast cells (CTMC) and mucosal mast cells (MMC). These populations differ in their location, cell size, staining characteristics, ultrastructure, mediator content and T-cell dependency¹⁻³. Several investigators⁴⁻⁹ recently reported a further subclass of mast cells which arise when normal mouse haematopoietic cells are cultured with interleukin-3 (IL-3); IL-3 is an activity similar or identical to mast-cell growth factor, histamine-producing factor, or P-cell stimulating factor⁴⁻⁹. These cultured mast cells are in many ways similar to MMC; they stain with Alcian blue but not safranin, contain chondroitin sulphate E proteoglycan rather than heparin proteoglycan and have relatively low histamine content, as do MMC¹⁰⁻¹³. Although proliferation of MMC is known to be T-cell dependent *in vivo* and thought to be IL-3-dependent *in vitro*, the factors on which CTMC proliferation depends remain elusive. Here we show that mature CTMC purified from mouse peritoneal cells can proliferate *in vitro* in methylcellulose culture and maintain the appearance and function of CTMC. We also present evidence that mature CTMC cannot proliferate in the presence of pure IL-3 alone.

When normal mouse peritoneal cells were cultured in methylcellulose medium containing pokeweed mitogen-stimulated spleen-cell-conditioned medium (PWMSCM), pure mast-cell colonies developed¹⁴. Detailed histochemical analyses of individual mast-cell colonies, together with observations of colony shape and size, allowed us to classify them into two types, type 1 and type 2 (Fig. 1*a,d*). The primary difference between colony types 1 and 2 was size: type 1 were smaller and more compact and revealed no signs of degeneration over 30 days of cultivation without additional feeding. Type 1 colonies contained 20 to 500 large mast cells with a single nucleus and many large granules (Fig. 1*b*); type 2 usually contained >2,000 small mast cells, with relatively small granules on day 16 of cultivation (Fig. 1*e*) and occasionally bilobed nuclei.

The two types of colonies could be more clearly distinguished by staining with Alcian blue-safranin or Berberine sulphate, or by measuring cellular histamine content. Cytocentrifuge preparations stained with Alcian blue-safranin and Berberine sulphate showed that most of the type 1 mast cells were both safranin-positive and Berberine sulphate-positive, whereas most of the type 2 mast cells stained with Alcian blue but not with safranin or Berberine sulphate (Table 1). Berberine sulphate staining has been reported to be highly specific for connective tissue-type mast cells, which contain heparin^{15,16}. We confirmed this by demonstrating that the staining was blocked by treating the cells with heparinase but not chondroitinase ABC. The histamine content of type 1 cells was significantly greater than that of type 2 cells (Table 1). These data strongly suggest that most type 1 cells are CTMC-like^{1-3,17}; by contrast, the type 2 cells resemble bone marrow-derived cultured mast cells, which have some characteristics in common with MMC^{3,10,11,17}.

We separated the precursors of type 1 mast-cell colonies from those of type 2 mast-cell colonies by removal of phagocytes followed by density gradient centrifugation (Table 2). Both types of precursor was enriched by the removal of phagocytes. Eight fractions were separated by Percoll density-gradient

Fig. 1 Mast-cell colonies. *a*, Typical type 1 mast-cell colony on day 16 of culture; *b*, May-Grunwald-Giemsa-stained smear on the type 1 mast-cell colony shown in *a*; *c*, Berberine sulphate staining of the same colony; most mast cells show fluorescence; *d*, typical type 2 mast-cell colony on day 16 of culture; *e*, May-Grunwald-Giemsa-stained smear of the type 2 mast-cell colony shown in *d*.

Methods. Peritoneal cells from 20-25-week-old BDF₁ mice were collected as described previously¹⁴. Methylcellulose cultures were carried out by using a modification of the technique described previously^{14,29}. Culture mixture (1 ml) containing 1×10^5 peritoneal cells, α -medium (Flow), 0.8% methylcellulose (Shinetsu Kagaku Co), 30% fetal bovine serum (Flow), 1% deionized bovine serum albumin (Sigma), 10^{-4} M mercaptoethanol (Eastman) and serum-free PWMSCM was incubated at 37 °C in a humidified atmosphere flushed with 5% CO₂ in air. Serum-free PWMSCM was prepared as described previously³⁰. Mast-cell colonies containing ≥ 20 cells were counted on day 16. Individual colonies were lifted from methylcellulose medium using 3 μ l Eppendorf pipette under a microscope and collected in Eppendorf microcentrifuge tubes containing 0.5 ml phosphate-buffered saline. After washing, the samples were spun in a cyto-centrifuge (Shandon) (800 r.p.m. for 5 min) and stained with May-Grunwald-Giemsa, Alcian blue-safranin³¹ or Berberine sulphate¹⁶.

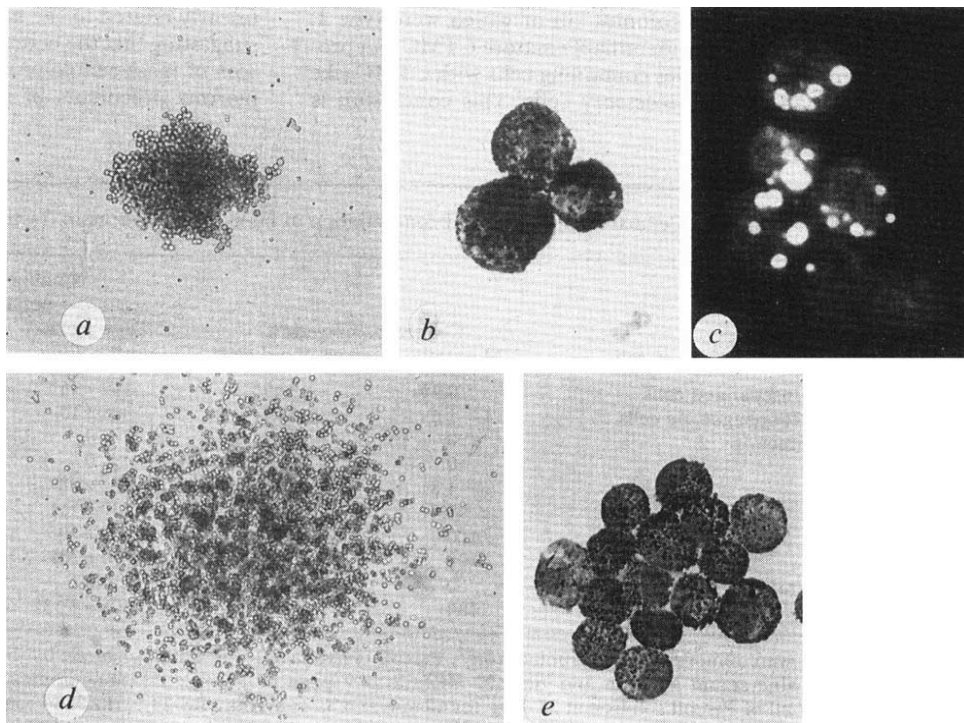


Table 1 Differences between types 1 and 2 mast-cell colonies

Type of mast-cell colony	No. of cells per colony	Cell size	Positive cells stained with			Histamine content (ng per 10 ⁵ cells)
			Berberine sulphate	Alcian blue	Safranin	
1	258 ± 35* (20–500)†	18.9 ± 1.9 µm (15–23)*	65.1 ± 21.2 (18–98)	77.4 ± 8.8 (55–87)	98.8 ± 0.9 (94–100)	64.74
2	8,500 ± 558 (2,000–18,500)	13.3 ± 3.7 (7–21)	0	100	1.4 ± 0.4 (0–5)	5.88

Colony size was determined by direct cell counting *in situ* with ×200 magnification on an inverted microscope when colonies consisted of less than 500 cells. When colonies contained >500 cells, a counting chamber was used. We analysed 50 types 1 and 2 mast-cell colonies. The histamine content was assayed using pooled colonies.

* Mean ± s.d.; † range.

centrifugation and individual fractions were analysed for cellular composition and relative concentration of the two types of precursor. The proportion of morphologically identifiable mast cells was increased in the heavier fractions and roughly paralleled the number of type 1 colonies. For example, fraction 7 ($\geq 1.13 \text{ g ml}^{-1}$) was an almost pure population of large mast cells all of which stained with both safranin and Berberine sulphate but not with Alcian blue, and were thus classed as mature CTMC¹⁸. When these cells were cultured in methylcellulose medium containing PWMSCM approximately half the cells gave type 1 colonies, but no type 2 colonies developed (experiment 2, Table 2). It is of interest that, although these colonies contained mostly safranin-positive, Berberine sulphate-positive cells, about half the colonies contained small numbers of Alcian blue-positive, safranin-negative cells which are thought to be a more primitive stage than CTMC¹⁸. The clonal origin of type 1 colonies was confirmed using single-cell culture¹⁹. Cells in fraction 7 were plated at a concentration of 1×10^3 per ml in methylcellulose medium containing PWMSCM in 35-mm plastic dishes. A single mast cell was identified morphologically, lifted from the medium using a fine Pasteur pipette attached to a Leitz micromanipulator, transferred to a fresh culture dish and incubated for 16 days. Fifteen of 50 transferred single mast cells produced colonies, all of which were type 1. Thus some morphologically identifiable mature CTMC can proliferate *in vitro* to give colonies containing cells with CTMC-like features in the absence of accessory cells. This conclusion is

supported in part by recent *in vivo* studies indicating that some peritoneal mast cells retain extensive proliferative potential after injection into the skin of mast-cell-deficient mice²⁰.

Sequential replating experiments of individual type 1 colonies derived from the cells in fraction 7 confirmed the proliferation potential of mature CTMC. Most of the primary colonies produced many secondary and tertiary colonies in methylcellulose medium containing PWMSCM. Although all the primary colonies contained mostly Berberine sulphate-positive, safranin-positive cells, there was a gradual decrease in the incidence of Berberine sulphate-positive mast cells and an increase in that of Alcian blue-positive, safranin-negative mast cells in individual secondary and tertiary colonies (data not shown). These findings suggest that some mature CTMC may be able to 'regress' and become a more immature type of CTMC or acquire some of the features of MMC. It will be necessary to perform a detailed biochemical examination of individual colonies to confirm this possibility.

Cells in fraction 2 ($\leq 1.05 \text{ g ml}^{-1}$; lymphocytes, lymphoblastoid cells and no detectable mast cells) produced only type 2 mast-cell colonies, consisting of Alcian blue-positive cells which did not stain with either safranin or Berberine sulphate (Table 2). The number of type 2 mast-cell colonies that developed was linearly related to the number of cells plated (data not shown), suggesting that the colonies were of clonal origin. Type 2 precursors of mast-cell colonies have a similar (low) density to bone marrow precursors of factor-dependent cultured mast cells²¹.

Table 2 Cellular composition and concentration of mast-cell precursors in fractionated peritoneal cells

Experiment	Origin of cells	Cellular composition		Number of colonies per 10 ⁵ cells		Total plating efficiency
		Mast cells	Lymphoid cells	Type 1 mast cell colony	Type 2 mast cell colony	
1	Unseparated cells	0.9%	53.9%	102 ± 25*	2 ± 1	0.1%
	Nonphagocytic cells	1.6	98.1	308 ± 15	4 ± 1	0.3
	Fraction 2	0	100	0	22 ± 2	0.02
	3	0	100	0	7 ± 2	0.01
	4	1.8	98.2	321 ± 28	3 ± 1	0.3
	5	97.9	2.1	8,230 ± 151	0	8.2
	6	100	0	19,930 ± 505	0	19.9
	7	100	0	16,210 ± 309	0	16.2
2	Fraction 2	0	100	0	12 ± 1	0.01
	7	100	0	50,608 ± 2,131	0	50.6

Phagocytic cells were removed using carbonyl iron³². Percoll (Pharmacia) was made isosmotic by the addition of 1 part of 10× Hanks' solution containing 1% bovine serum albumin and 200 mM HEPES to 9 parts of Percoll. Continuous density gradients of Percoll were generated by centrifugation of 7 ml of Percoll solution at 23,000g for 60 min at 4 °C in a Tomy RS-20 FIIIS fixed rotor. Cell suspension (1 ml) containing 1×10^7 nonphagocytic cells was layered on the top of Percoll continuous gradients and centrifuged at 500g for 30 min²⁰. Although eight fractions were separated, fractions 1 and 8 were discarded because the former contained degenerated cells and the latter contained a small number of macrophages which phagocytosed iron particles. Cellular composition was determined on cytocentrifuge preparations stained with May-Grunwald-Giemsa. Each fraction was cultured before (Fig. 1 legend).

* Mean ± s.d. of triplicated cultures.

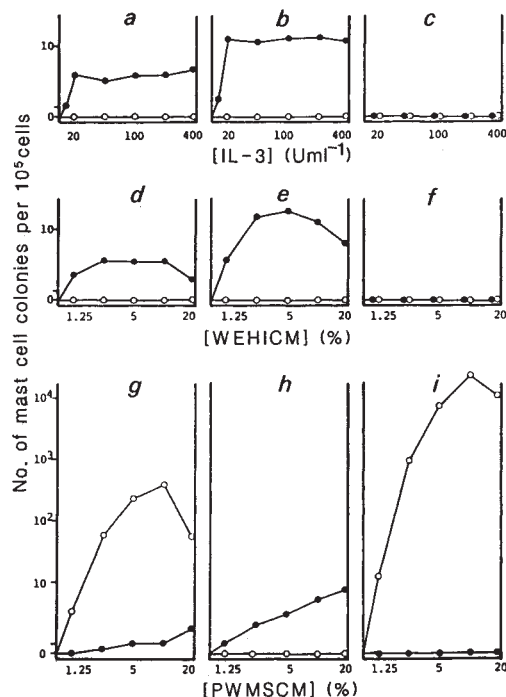


Fig. 2 Dependence of colony growth on growth factors. Growth of colony types 1 and 2 from *a, d, g*, unseparated peritoneal cells; *b, e, h*, cells in fraction 2; *c, f, i*, cells in fraction 7 at varying concentrations of *a-c*, pure IL-3 in culture; *d-f*, WEHICM; and *g-i*, PWMSCM. Open circles, type 1 mast-cell colony; filled circles, type 2 colony.

Thus the precursors for two distinct types of mast cell can be separated by Percoll density-gradient centrifugation.

We investigated the growth factor requirements of the two types of mast-cell precursors. Homogeneous populations of mast cells resembling MMC are obtained *in vitro* from mouse bone marrow, spleen, fetal liver and lymphoid tissue in cultures supplemented with conditioned medium from stimulated T cells or the WEHI-3B cell line⁴⁻⁷. It has been demonstrated that the biologically active component in the conditioned medium is IL-3^{8,9,12,22}. However, little is known about the factor(s) supporting CTMC proliferation *in vitro*. Complex culture systems have been reported in which heparin-containing mast cells differentiate when the culture contains embryonic fibroblasts and lymphocytes from mice preimmunized to horse serum²³. Recent

observations show that rat CTMC can be maintained *in vitro* in the presence of living fibroblasts but do not divide, and that fibroblast-derived conditioned medium does not have any ability to maintain CTMC²⁴. We next examined whether conditioned media (other than PWMSCM) or pure IL-3 support the proliferation of either type 1 or type 2 colonies in our methylcellulose culture system. Neither type of colony will grow in the presence of lung-conditioned medium (containing granulocyte-macrophage colony stimulating factor and granulocyte-colony stimulating factor^{25,26}). L-cell conditioned medium (containing macrophage colony stimulating factor²⁷) also failed to stimulate colony growth.

Figure 2 shows the plating efficiencies of two types of mast-cell colonies from unfractionated or fractionated peritoneal cells stimulated by different concentrations of pure IL-3, WEHI-3B conditioned medium (WEHICM) or PWMSCM. No type 1 mast-cell colonies were observed in the presence of up to 400 U ml⁻¹ of pure IL-3 in the culture. WEHICM also failed to stimulate type 1 mast-cell colony growth. PWMSCM stimulated type 1 mast-cell colony growth from unfractionated cells and fraction 7 significantly at a maximum concentration of 10%. Pure IL-3 produced type 2 mast-cell colonies from unfractionated cells and fraction 2 in a dose-dependent fashion, reaching a plateau at 20 U ml⁻¹ IL-3. Pure IL-3 also supported the growth of secondary mast-cell colonies consisting of cells which stained with Alcian blue but not with safranin or Berberine sulphate upon replating of type 2 colonies. Our experiments show that the two types of mast-cell precursors are different in origin and in requirement for growth factors. The IL-3 dependent growth of the low density precursors for type 2 colonies is consistent with a previous report that bone marrow precursors are distinct from mast cells, have a low density and proliferate in response to IL-3²¹. Our PWMSCM contained some factor(s) that stimulate mature CTMC proliferation and is not IL-3. However, our data do not exclude the possibility that IL-3 can act as a co-stimulator for CTMC proliferation, because there was a small amount of IL-3 in the cultures containing 10% (v/v) PWMSCM (3.6 U ml⁻¹: 100 U activity was arbitrarily defined as that amount of IL-3 per ml that gives 30% of maximal ³H-thymidine incorporation into the IL-3 dependent cell line 32 Dcl²⁸). Further studies are needed to determine the nature and source of the factor(s) responsible for mature CTMC proliferation.

The CTMC culture system described here should allow the elucidation of the cellular and molecular mechanisms that control the proliferation of mature CTMC. An examination of the changes in phenotype and processes involved in turnover of mediators during extensive proliferation should also be possible.

We thank Drs Y. Kitamura and A. Komiyama for helpful advice and Dr J. N. Ihle for pure IL-3. Supported by grants from the Ministry of Education, Science and Culture, Japan.

Received 5 March; accepted 6 August 1986.

1. Enerback, L. *Monogr. Allerg.* **17**, 222-232 (1981).
2. Bienenstock, J., Befus, A. D., Pearce, F., Denburg, J. & Goodacre, R. *Allerg. clin. Immun.* **70**, 407-412 (1982).
3. Galli, S. J., Dvorak, A. M. & Dvorak, H. F. *Prog. Allerg.* **34**, 1-141 (1984).
4. Nabel, G., Galli, S. J., Dvorak, A. M., Dvorak, H. F. & Cantor, H. *Nature* **291**, 332-334 (1981).
5. Schrader, J. W., Lewis, S. J., Clarke-Lewis, I. & Culvenor, J. G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 323-327 (1981).
6. Razin, E., Cordon-Cardo, C. & Good, R. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2559-2561 (1981).
7. Yung, Y. P., Eger, R., Tertian, G. & Moore, M. A. S. *J. Immun.* **127**, 794-799 (1981).
8. Ihle, J. N. *et al. J. Immun.* **131**, 282-287 (1983).
9. Yokota, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 1070-1074 (1984).
10. Haig, D. M., McKee, T. A., Jarrett, E. E., Woodbury, R. & Miller, H. R. P. *Nature* **300**, 188-190 (1982).
11. Razin, E., Stevens, R. L., Akiyama, F., Schmidt, K. & Austen, K. J. *biol. Chem.* **257**, 7229-7236 (1982).
12. Razin, E., *et al. J. Immun.* **132**, 1479-1486 (1984).
13. Galli, S. J. *et al. J. Cell Biol.* **95**, 435-444 (1982).
14. Kobayashi, T. *et al. J. Immun.* **136**, 1378-1384 (1986).
15. Enerback, L. *Histochemistry* **42**, 301-313 (1974).
16. Dimlich, R. V. M., Meineke, H. A., Reilly, F. D. & McCuskey, R. S. *Stain Tech.* **55**, 217-223 (1980).
17. Nakano, T. *et al. J. exp. Med.* **162**, 1025-1043 (1985).
18. Pretlow, T. G. & Gassady, I. M. *Am. J. Path.* **61**, 323-338 (1970).
19. Nakahata, T. *et al. Blood* **65**, 1010-1016 (1985).
20. Sonoda, T. *et al. J. exp. Med.* **160**, 138-151 (1984).
21. Yung, Y. P., Wang, S. Y. & Moore, M. A. S. *J. Immun.* **130**, 2843-2848 (1983).
22. Rennick, D. M. *et al. J. Immun.* **134**, 910-914 (1985).
23. Bland, C. E., Ginsburg, H., Silbert, J. E. & Metcalfe, D. D. *J. biol. Chem.* **257**, 8661-8666 (1982).
24. Levi-Schaffer, F. *et al. J. Immun.* **135**, 3454-3462 (1985).
25. Burgess, A. W., Camakaris, J. & Metcalf, D. *J. biol. Chem.* **252**, 1998-2003 (1977).
26. Nicola, N. A., Metcalf, D., Matsumoto, M. & Johnson, G. R. *J. biol. Chem.* **258**, 9017-9023 (1983).
27. Stanley, E. R. & Heard, B. M. *J. biol. Chem.* **252**, 4305-4312 (1977).
28. Frei, K., Bodmer, S., Schwerdel, C. & Fontana, A. *J. Immun.* **135**, 4044-4047 (1985).
29. Nakahata, T., Spicer, S. S., Cantey, J. R. & Ogawa, M. *Blood* **60**, 352-361 (1982).
30. Nakahata, T. & Ogawa, M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3843-3847 (1982).
31. Spicer, S. S. *J. Histochem. Cytochem.* **8**, 18-35 (1960).
32. Kubota, K., Mizoguchi, H., Miura, Y., Suda, T. & Takaku, F. *Expl. Hemat.* **8**, 339-344 (1980).

A chimaeric receptor allows insulin to stimulate tyrosine kinase activity of epidermal growth factor receptor

Heimo Riedel, Thomas J. Dull, Joseph Schlessinger* & Axel Ullrich

Department of Developmental Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

The cell surface receptors for insulin and epidermal growth factor (EGF) appear to share a common evolutionary origin, as suggested by structural similarity of cysteine-rich regions in their extracellular domains and a highly conserved tyrosine-specific protein kinase domain¹⁻³. Only minor similarity is found outside this catalytic domain, as expected for receptors that have different ligand specificities and generate different biological signals^{4,5}. The EGF receptor is a single polypeptide chain⁶ but the insulin receptor consists of distinct α and β subunits⁷ that function as an $\alpha_2\beta_2$ heterotetrameric receptor complex⁸⁻¹¹. Provoked by this major structural difference in two receptors that carry out parallel functions, we have designed a chimaeric receptor molecule comprising the extracellular portion of the insulin receptor joined to the transmembrane and intracellular domains of the EGF receptor to investigate whether one ligand will activate the tyrosine kinase domain of the receptor for the other ligand. We show here that the EGF receptor kinase domain of the chimaeric protein, expressed transiently in simian cells, is activated by insulin binding. This strongly suggests that insulin and EGF receptors employ closely related or identical mechanisms for signal transduction across the plasma membrane.

Cloned human EGF receptor¹ and insulin receptor complementary DNAs^{2,3} were used in combination with simian virus 40 (SV40) early promoter transcriptional elements (Fig. 1a) to generate expression constructs coding for the complete human insulin receptor (HIR) and a chimaeric receptor (IER), shown in Fig. 1b. The insulin receptor is normally made as a precursor and proteolytically processed to generate the α and β subunits⁴. IER was generated by fusing the coding region of the signal sequence and extracellular portion of the insulin receptor precursor (amino acids -27-914) to the coding sequence of the transmembrane and cytoplasmic domains of the EGF receptor¹ (amino acids 622-1,186). Using this chimaeric construct we asked whether the hybrid receptor would be proteolytically processed into insulin receptor α subunit and the recombinant I β E subunit, whether a functional complex would be formed and directed to the plasma membrane and whether insulin binding could stimulate autophosphorylation of the EGF receptor cytoplasmic domain in the chimaeric cell surface membrane glycoprotein.

Expression of the complete insulin receptor resulted in protein products of relative molecular mass (M_r) 95,000 (95 K) and 180 K and a diffuse band around 135 K, as determined by ³⁵S-methionine metabolic labelling (Fig. 2a). These polypeptides represent the β -subunit, the uncleaved precursor and the glycosylated α subunit of the insulin receptor^{2,9,12-14}. Proteolytic processing of the native insulin receptor precursor is inefficient under transient expression conditions in COS-7 cells, as one can see from the amount of HIR precursor remaining.

By analogy with the native receptors, COS-7 cells expressing the IER receptor chimaera would be expected to synthesize a 135 K glycosylated α -chain, a comparably sized 120 K glycopolypeptide composed of the intracellular and transmem-

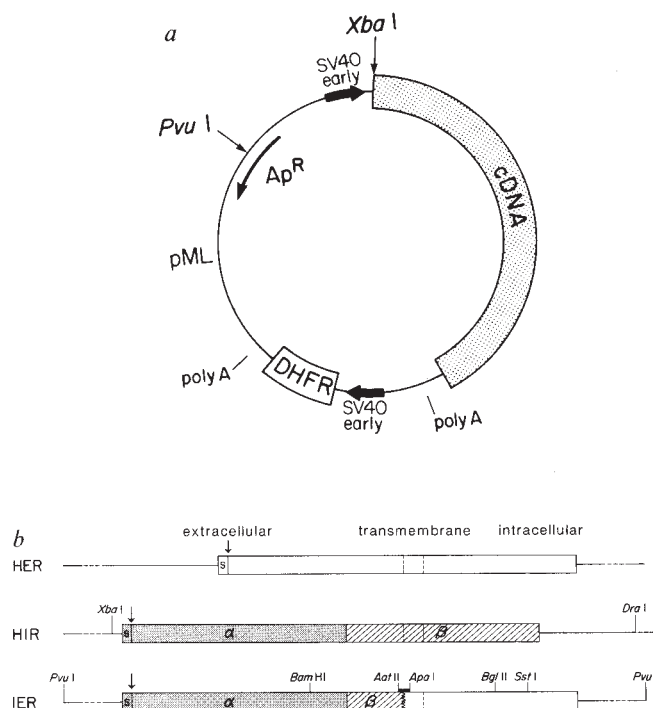


Fig. 1 *a*, Receptor expression plasmid; *b*, schematic comparison of normal and chimaeric receptor constructs. *a*, The region encoding different receptor constructs is shaded (cDNA). Early SV40 promoter sequences are shown by heavy black arrows and poly(A) addition sites are indicated. The dihydrofolate reductase (DHFR) coding sequence¹⁹ and restriction sites used for plasmid constructions are shown. Expression of both cDNAs was controlled by promoter sequences of the SV 40 early region. Sequences of the *Escherichia coli* plasmid pML (ref. 20) were present to allow plasmid replication in *E. coli*. *b*, Coding regions of human EGF receptor (HER), human insulin receptor (HIR) and insulin-EGF receptor chimaera (IER). Shaded regions, HIR α subunit sequences; striped regions, HIR β sequences; unshaded regions, HER sequences. The sequences are aligned at their transmembrane domains (not to scale). S, signal sequences; arrows, signal sequence cleavage sites; vertical lines, receptor precursor cleavage sites; zigzag line, junction between heterologous receptor cDNAs; black bar, synthetic oligonucleotide used for the connection.

Methods. Construction of the HIR expression plasmid: a *SalI* restriction fragment containing the entire HIR protein coding sequence² was introduced into the *SalI* site in the polylinker region of the plasmid pUC12. DNA was amplified and molecules containing the 5' end of the HIR coding sequence next to the pUC12 *XbaI* site were used to generate a *XbaI*-*DraI* restriction fragment containing the complete HIR protein coding sequence. This fragment was introduced into the mammalian expression vector (Fig. 1a) at an *XbaI* and a filled-in *BamHI* site in the orientation *XbaI*/*XbaI*-*DraI*/*BamHI*. Construction of the IER expression plasmid: a 931 bp *BamHI*-*AatII* restriction fragment from the HIR expression plasmid and a 1,150 bp *ApaI*-*SstI* restriction fragment of the human EGF receptor sequence¹ were joined at their respective *AatII* and *ApaI* sites with the complementary synthetic oligonucleotides 5'-CCCGTCAAATATCGCCACTGGGATGGTGGGGGCC-3' and 5'-CCCACCATCCCAGTGGCGATATTTGACGGGACGT-3'. The ligated fragments were inserted into a *SstI*/*BamHI* digested pUC12 vector, giving a combination of the sequences coding for the extracellular domain of the insulin receptor with the transmembrane and cytoplasmic domains of the EGF receptor. A 965 bp *BamHI*-*ApaI* fragment of this plasmid, containing the IER junction, was then ligated to a 3.1 kb *PvuI*-*BamHI* fragment of the HIR expression plasmid, which contained the rest of the HIR coding sequence and parts of the expression vector, and to two fragments (810 bp *ApaI*-*BglII* and 5.3 kb *BglII*-*PvuI*) coding for the transmembrane and cytoplasmic domains of the human EGF receptor, with the latter also containing sequences of the expression vector.

* Biotechnology Research Center, Meloy Labs, 4 Research Court, Rockville, Maryland 20850, USA.

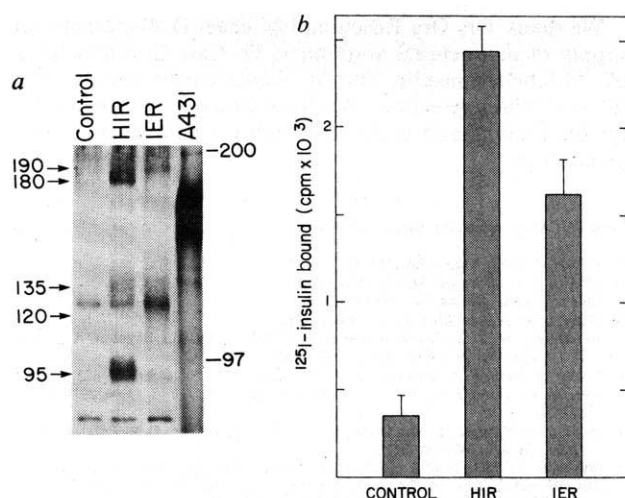


Fig. 2 *a*, ^{35}S -methionine labelled cDNA expression products; *b*, ^{125}I -insulin binding by chimaeric receptors. *a*, COS-7 cells were transfected with the constructs indicated (10 μg supercoiled DNA) or with a control expression vector lacking the receptor cDNA insert²¹. A431 epidermoid carcinoma cells were used for EGF receptor control experiments. COS-7 cells 44 h after transfection were washed twice with methionine-free medium prepared from a MEM select amine kit (Gibco) and incubated for 9 h with 200 μCi ^{35}S -methionine at 800 Ci mmol^{-1} per 8 cm dish. Proteins were immunoprecipitated with monoclonal antibodies to human insulin receptor²² (MAb CII25.3) or human EGF receptor²³ (MAb RI), as described²⁴, and were analysed on 7.5% reducing SDS polyacrylamide gels. Sizes of polypeptide markers are indicated (myosin H-chain, 200 K; phosphorylase B, 97.4 K). Arrows, immunoprecipitated expression products. *b*, Transfected COS-7 cells were washed twice with phosphate-buffered saline (PBS) after 53 h and incubated in 1 ml per 2.2 cm well of serum-free cell culture medium containing 0.2% bovine serum albumin (Sigma), bacitracin (0.5 mg ml^{-1}) (Sigma) and ^{125}I -labelled insulin (0.5 μCi per well) at 93 μCi μg^{-1} (provided by Dr Carl Grunfeld, UCSF) for 2 h at 21 $^{\circ}\text{C}$. Cells were washed 3 times with PBS at 4 $^{\circ}\text{C}$, lysed in 0.5 ml 0.1% SDS, 0.1 M NaOH for 30 min at 37 $^{\circ}\text{C}$, and the total radioactivity was measured. Average of three determinations shown.

brane sequences of the EGF receptor fused to the insulin receptor β subunit extracellular portion (I β E), and, if cleavage of IER proreceptor sequences is inefficient, a 190 K polypeptide. Each of these products was detected in cells expressing the IER chimaera; the I β E fusion glycoprotein subunit migrates as a diffuse band just below the 135 K α -subunit polypeptide (Fig. 2a; see also Fig. 3).

Binding experiments with ^{125}I -labelled insulin (Fig. 2b) indicated that the products of the HIR and IER constructs were expressed in functional form on the surface of COS cells. Insulin binding was highest for the HIR construct and the amount of binding detected was proportional to the levels of the glycopolypeptides present in the transfected cells (Fig. 2a).

To test their intrinsic kinase activity, the expressed polypeptides were immunoprecipitated from lysates of transfected COS-7 cells using antibodies against the EGF receptor (Fig. 3a) or the insulin receptor (Fig. 3b,c), and [γ - ^{32}P]ATP was added. Insulin-stimulated autophosphorylation of the human insulin receptor β subunit was clearly detected above the endogenous COS-7 control, at substrate concentrations (100 μM ATP) required by the insulin receptor kinase^{15,16} that result in reduced specific activity of the radioactive tracer ([γ - ^{32}P]ATP) and thus a weak signal (HIR, Fig. 3b). In contrast, only subpicomolar (0.1 pM) concentrations of ATP were needed to detect EGF-stimulated EGF-receptor kinase activity¹⁷ (Fig. 3a,c), conditions under which the insulin receptor kinase is inactive.

As shown in Fig. 3a and c, the IER chimaeric receptor displays insulin-stimulated autophosphorylation activity. Insulin stimu-

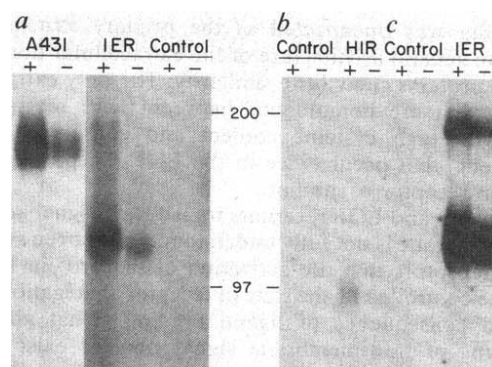


Fig. 3 Autophosphorylation of normal and chimaeric receptors. Immunoprecipitation from: *a*, A431 cells using human EGF receptor mouse monoclonal antibody RI²³, or from COS-7 cells transfected with the IER construct or with a control expression vector lacking a receptor cDNA insert, using human EGF receptor rabbit polyclonal antibody RK-2 (ref. 24); *b*, HIR-transfected or control cells using human insulin receptor mouse monoclonal antibody CII25.3; *c*, IER-transfected or control cells using CII25.3. Detergent cell lysates were stimulated with insulin (+) (EGF for A431 cells) or mock-stimulated with PBS (-).

Methods. Insulin or EGF stimulation of autophosphorylation was induced by incubating the cell lysates before immunoprecipitation with either 200 nM insulin (Sigma) for 1 h or 1 μM EGF for 15 min at 4 $^{\circ}\text{C}$. Antibody RI, RK-2 or CII25.3 prebound for 30 min to protein A-Sepharose was added, and incubated for 15 min at 4 $^{\circ}\text{C}$. The washed immunoprecipitates, in a volume of 30 μl , were adjusted to 5 mM MnCl_2 , and 15 μCi of [γ - ^{32}P] ATP was added for 1 min at 4 $^{\circ}\text{C}$. The final ATP concentration was 0.1 pM (*a* and *c*) or 100 μM (*b*). The reaction was stopped by adding 20 μl of 3 times concentrated SDS sample buffer, and samples were analysed on 5%/7% SDS polyacrylamide gels²⁵.

lation of chimaeric protein autophosphorylation was observed in immunoprecipitates obtained using antibodies to either the insulin receptor extracellular domain (Fig. 3c) or the EGF receptor intracellular domain (Fig. 3a). This experiment rules out the possibility that small amounts of insulin-stimulated, endogenous COS cell insulin receptor catalysed phosphorylation of the chimaeric receptors in the immunoprecipitates. The chimaeric IER kinase was active at low ATP concentrations (0.1 pM) consistent with the enzymatic characteristics of the EGF receptor kinase domain. Interestingly, phosphorylation of the bands representing the I β E fusion glycoprotein subunit (120 K) and the uncleaved chimaeric IER proreceptor (190 K) was also induced by insulin. Like the native EGF receptor¹⁷, a maximal ligand induction was observed after 30 seconds at 4 $^{\circ}\text{C}$ (not shown). Thus the EGF receptor tyrosine kinase retains its original enzymatic characteristics when controlled by the insulin binding domain of a chimaeric receptor molecule.

We have described here the expression and functional characterization of a chimaeric receptor molecule composed of the extracellular insulin receptor sequences linked to the EGF receptor transmembrane and cytoplasmic domains. We have shown that binding of insulin to this IER chimaeric receptor molecule induces tyrosine kinase activity in the EGF receptor-derived intracellular domain. The observed tyrosine kinase activity is characteristic of the EGF receptor in that it was measurable at subpicomolar ATP concentrations, which do not activate the insulin receptor phosphotransferase^{15,16}. Therefore, the IER chimaeric receptor precursor must have been correctly processed and folded into a conformation permitting communication between two entirely heterologous receptor domains, at the same time preserving the kinetic properties of the EGF receptor tyrosine kinase.

These results directly demonstrate that insulin and EGF receptors use a common transmembrane signal transfer mechanism.

This finding was unexpected as the primary structure and, indeed, the general architecture of the extracellular domains of these two proteins show little similarity. The only extracellular structures that are homologous between these receptors are regions with high cysteine content and conserved cysteine arrangement² that occur twice in the EGF receptor and once per insulin receptor α subunit.

How insulin and EGF receptors transduce a signal across the plasma membrane is not fully understood. It has been suggested for both receptors that the activation of the protein kinase is likely to be controlled by the state of receptor aggregation, which may be a consequence of ligand binding. What alternative mechanisms of transmembrane signal transfer exist for this receptor type? Domain communication through a transmembrane α helix is difficult to envisage, as is interaction of the homologous, cysteine-rich units with a common, heterologous membrane component needed for signal transfer. We favour the notion that a ligand-induced conformational change in the receptor extracellular domain triggers receptor dimerization resulting in activation of intracellular enzymatic activities.

Whereas the EGF receptor may have to self-associate to induce kinase activity on the cytoplasmic side of the membrane, insulin and IGF-1 receptors¹⁸, both disulphide-linked heterotetramers, may be structured in this way to avoid the need for aggregation and predispose these receptors to respond to ligand. The ability of the chimaeric receptors to transduce a signal suggests that the structural differences between EGF and insulin receptors are variations on a common mechanism for signal transfer. These results open new avenues of investigation into the mechanisms of transmembrane signalling and receptor function.

We thank Drs Ora Rosen and Michael D. Waterfield for the supply of monoclonal antibodies, Dr Carl Grunfeld for a gift of ¹²⁵I-labelled insulin, and M. Vasser and P. Jhurani for synthesis of oligonucleotides. We are indebted to Dr Suzanne Pfeffer for discussions and to Jeanne Arch for help in preparing the manuscript.

Received 1 May; accepted 8 September 1986.

1. Ullrich, A. *et al. Nature* **309**, 418-425 (1984).
2. Ullrich, A. *et al. Nature* **313**, 756-761 (1985).
3. Ebina, Y. *et al. Cell* **40**, 747-758 (1985).
4. Kahn, C. R. *et al. Rev. Med.* **36**, 429-451 (1985).
5. Herschman, H. R. in *Control of Animal Cell Proliferation* (eds Boynton, A. L. & Leffert, H. L.) 169-199 (Academic, New York, 1985).
6. Cohen, S., Fava, R. A. & Sawyer, S. T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6237-6241 (1982).
7. Kasuga, M., Hedo, J. A., Yamada, K. M. & Kahn, C. R. *J. biol. Chem.* **257**, 10392-10399 (1982).
8. Hedo, J. A., Kasuga, M., Van Obberghen, E., Roth, J. & Kahn, C. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4791-4795 (1981).
9. Massague, J., Pilch, P. F. & Czech, M. P. *J. biol. Chem.* **256**, 3182-3190 (1981).
10. Van Obberghen, E. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 1052-1056 (1981).
11. Fujita-Yamaguchi, Y. *J. biol. Chem.* **259**, 1206-1211 (1984).
12. Kasuga, M., Kahn, C. R., Hedo, J. A., Van Obberghen, E. & Yamada, K. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6917-6921 (1981).
13. Reed, B. C., Ronnett, G. V. & Lane, D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2908-2912 (1981).
14. Deutsch, P. J., Wan, C. F., Rosen, O. M. & Rubin, C. S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 133-136 (1983).
15. White, M. F., Haring, H.-U., Kasuga, M. & Kahn, C. R. *J. biol. Chem.* **259**, 255-264 (1984).
16. Petruzzelli, L., Herrera, R. & Rosen, O. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3327-3331 (1984).
17. Erneux, C., Cohen, S. & Garbers, D. *J. biol. Chem.* **258**, 4137-4142 (1983).
18. Massague, J. & Czech, M. P. *J. biol. Chem.* **257**, 5038-5045 (1982).
19. Simonsen, C. C. & Levinson, A. D. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2495-2499 (1983).
20. Lusk, M. & Botchan, M. *Nature* **293**, 79-81 (1981).
21. Graham, F. L. & van der Eb, A. J. *Virology* **52**, 456-467 (1973).
22. Ganguly, S., Petruzzelli, L. M., Herrera, R., Stadtmayer, L. & Rosen, O. M. *Curr. Top. cell Regul.* **27**, 83-94 (1985).
23. Waterfield, M. D. *et al. J. cell Biochem.* **20**, 149-161 (1982).
24. Kris, R. M. *et al. Cell* **40**, 619-625 (1985).
25. Laemmli, U. K. *Nature* **227**, 680-685 (1970).

Cloning of complementary DNA encoding T-cell replacing factor and identity with B-cell growth factor II

Tatsuo Kinashi, Nobuyuki Harada*, Eva Severinson†, Toshizumi Tanabe, Paschalis Sideras†, Mikio Konishi, Chihiro Azuma, Akira Tominaga*, Susanne Bergstedt-Lindqvist†, Masazumi Takahashi, Fumihiko Matsuda, Yoshio Yaoita, Kiyoshi Takatsu* & Tasuku Honjo

Department of Medical Chemistry, Kyoto University Faculty of Medicine, Yoshida, Sakyo-ku, Kyoto 606, Japan

* Department of Biology, Institute for Medical Immunology, Kumamoto University Medical School, Kumamoto 860, Japan

† Department of Immunology, Stockholm University, S-106 91 Stockholm, Sweden

Proliferation and maturation of antigen-stimulated B cells are regulated by several soluble factors derived from macrophages and T cells^{1,2}. These soluble factors are functionally divided into two groups: B-cell growth factor (BCGF), thought to be involved in B-cell proliferation; and B-cell differentiation factor (BCDF), responsible for maturation of activated B cells into immunoglobulin-secreting cells³⁻⁶. This classification needs to be re-examined in the light of the recent cloning of complementary DNA encoding IgG₁ induction factor (interleukin-4, IL-4) from the 2.19 mouse T-cell line⁷. Recombinant IL-4 has BCGF and BCDF activities and affects B cells, T cells and mast cells (refs 7, 8; our unpublished data). Another well-characterized B-cell factor is T-cell replacing factor (TRF)⁹⁻¹², which, when secreted by the murine T-cell hybridoma B151K12, is defined by two activities¹⁰⁻¹²: induction of IgM secretion by BCL₁ leukaemic B-cell line; and induction of secondary anti-dinitrophenol (DNP) immunoglobulin G (IgG) synthesis *in vitro* by DNP-primed B

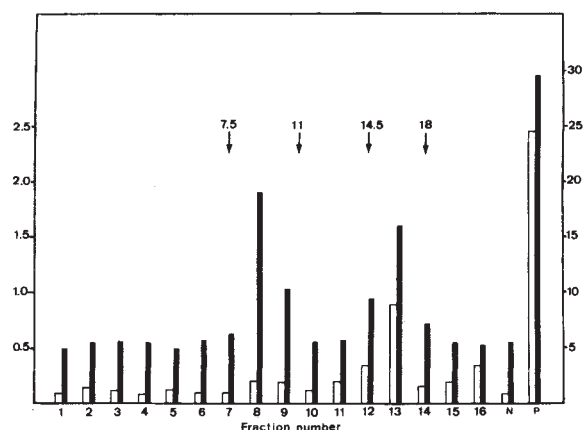


Fig. 1 Effects on BCL₁ cells of translational products directed by size-fractionated mRNA of 2.19 T-cell line. Poly(A)⁺ RNA was obtained from the Concanavalin A-activated 2.19 T-cell line and fractionated by centrifugation in a 5-22% sucrose gradient at 36,000 rpm for 15 h⁷. An aliquot of each of the 16 fractions was injected into *Xenopus* oocytes. Incubation media were collected after incubation for 36 h at 20 °C and added to the assay medium at the concentration of 1%. Right-hand scale, ³H-thymidine incorporation (closed bars; c.p.m. per culture $\times 10^{-3}$). Left-hand scale, anti-IgM plaque-forming response (open bars; plaque-forming cells per culture $\times 10^{-3}$ of BCL₁ cells, measured as described in Table 1. P, positive control where culture supernatants of B151 hybridoma were added to the medium at the concentration of 50%; N, negative control where phosphate buffer was injected into *Xenopus* oocytes instead of mRNA. Arrows, sizes of mRNA (in Svedberg units).

cells. Although TRF from B151K12 was classified as BCDF, purified TRF has BCGF-II activity¹³. To elucidate the molecular properties of TRF we isolated cDNA encoding TRF from the 2.19 T-cell line and report here the structure and multiple activities of this lymphokine.

Table 1 TRF and BCGF-II activities on BCL₁ of supernatants of cell culture and mRNA-injected oocytes

Cells of supernatants or templates of mRNA injected into oocytes	Supernatants added (%)	IgM plaque-forming cells per culture	Thymidine incorporation (c.p.m. per culture)
B151 hybridoma	25	1,834	7,069
	10	903	4,222
2.19 T-cell line	1.0	3,134	13,726
	0.25	1,716	6,847
HeLa cell	10	279	705
None	0	211	1,289
rIL-4 (pSP6KmIL4-374)	0.25	32	17,396
B151 TRF	25	1,159	21,242
None	0	60	4,822
<i>SacI</i> -digested library	5	500	8,396
<i>SalI</i> -digested library	5	799	14,342
Phosphate buffer	5	563	6,732
cDNA clone 16	5	199	1,728
18	5	233	1,969
23	5	3,132	17,970
27	5	366	2,001
42	5	215	1,650
51	5	201	1,862
53	5	180	1,883
none	5	150	1,654
<i>SalI</i> -digested clone-23	1	3,229	10,824
	0.25	1,662	6,615
	0.06	772	3,037
	0.015	232	1,625
None	0	211	1,289
<i>AccI</i> -digested clone-23	10	2,343	18,772
<i>SalI</i> -digested clone-23	10	2,277	23,386
None	0	208	4,904

Supernatants of cells or oocytes injected with mRNA were added to the culture medium of BCL₁ cells at indicated concentrations. BCL₁ cells (1.5×10^5) of an *in vivo* line were cultured in 0.2 ml RITC medium (serum-free, 5 mg ml⁻¹ bovine serum albumin, 50 μ M 2-mercaptoethanol and antibiotics) in addition to testing supernatants. After a 2-day culture, the protein A plaque assay¹⁹ was used to determine the numbers of IgM-secreting cells. BCL₁ proliferation was determined by ³H-thymidine uptake for the last 6 h of a 2-day culture. Culture medium of various cell lines were added to the assay medium at indicated concentrations. rIL-4, Supernatants of *Xenopus* oocytes injected with mRNA derived from pSP6KmIL4-374 as described previously⁷. B151 TRF prepared by partial purification from supernatants of B151 cells¹³. The total cDNA library consisted of 5×10^4 clones and the library DNAs digested with *SacI* or *SalI*. The detailed methods of mRNA synthesis by SP6 RNA polymerase and its injection to *Xenopus* oocytes were described previously⁷. Individual cDNA clones with 1-kb longer inserts containing *SacI* sites were digested with *SalI* and their mRNAs tested. Incubation medium of oocytes injected with distilled water was used as a negative control. The construct pSP6K-mTRF23 was digested either by *SalI* or by *AccI* and used as template for *in vitro* mRNA synthesis. All values are means in the triplicate experiments and standard errors of triplicate experiments ranged from 1.01 to 1.12.

Supernatants of the 2.19 T-cell line¹⁴ contain the TRF activity (Table 1) and, like the B151 supernatants, have the BCGF-II activity as assayed by induction of proliferation of BCL₁ cells. Poly(A)⁺RNA of 2.19 T cells was fractionated by sucrose density gradient centrifugation and an aliquot of each fraction was injected into *Xenopus* oocytes to test the presence of messenger RNA encoding BCGF-II and TRF by measuring proliferation stimulation and IgM plaque formation of BCL₁ cells, respectively. BCL₁ proliferation occurs in translation products directed by 9 and 16S mRNAs (Fig. 1). In contrast, IgM plaque-forming activity is found only in products of 16S mRNA. The 9S mRNA corresponds to the size of mRNA for IgG₁ induction factor (IL-4)⁷. Recombinant IL-4 produced by *Xenopus* oocytes shows BCL₁ proliferation activity but not IgM plaque-forming activity (Table 1). Furthermore, translation products directed by 15–18S mRNA of B151K12 cells exert TRF activity¹⁵. We therefore conclude that 16S mRNA of 2.19 T cells encodes another factor that has TRF activity. The same fraction of mRNA also encodes BCGF-II defined by the growth stimulation of BCL₁. The results also indicate that the BCL₁ proliferation assay does not discriminate between BSF-I (IL-4) and BCGF-II.

To clone cDNA encoding TRF we used the pSP6K cDNA library of the total poly(A)⁺ RNA of 2.19 T cells, the construction of which was described previously⁷. A DNA mixture of 5×10^4 independent cDNA clones was cleaved with *SacI* or *SalI* to linearize plasmid DNA and mRNAs were synthesized using SP6 RNA polymerase. The synthesized RNAs were injected into

Table 2 Stimulation of anti-DNP IgG plaque formation by rTRF

Pretreatment with anti-Thy 1.2	Antigen	TRF in culture (%)	Anti-DNP IgG plaque-forming cells per culture
—	None	None	127
—	DNP-KLH	None	1,792
+	None	None	127
+	DNP-KLH	None	199
+	DNP-OA	None	92
+	DNP-OA	rTRF 1	695
+	DNP-OA	rTRF 0.5	884
+	DNP-OA	rTRF 0.25	555
+	DNP-OA	rTRF 0.125	382
+	DNP-OA	B151 10	720
+	DNP-OA	B151 5	636
+	DNP-OA	B151 2.5	485
+	DNP-OA	B151 0.25	325

BALB/c mice were immunized with DNP-KLH and spleen cells isolated 6–8 weeks after the immunization and treated twice with anti-Thy 1.2 and complements. The cells (6×10^5) were cultured in 0.2 ml culture media (RPMI 1640) containing 10% fetal calf serum (FCS) together with 10 ng DNP-ovalbumin (DNP-OA) and TRF for 5 days. Anti-DNP IgG plaque-forming cells were measured as described¹¹. B151 indicates partially purified TRF from supernatants of B151 cells as described¹³.

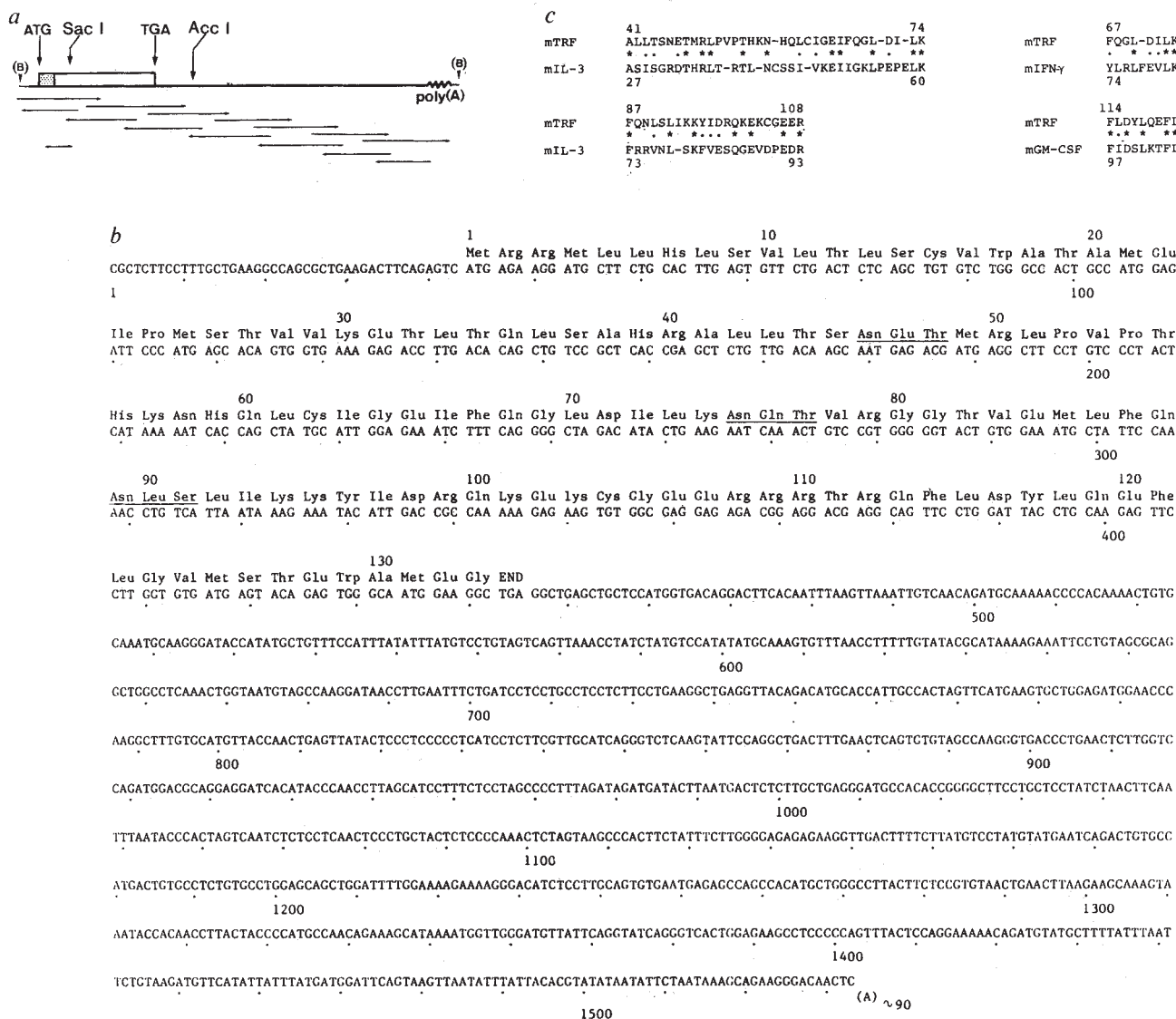


Fig. 2 *a*, Restriction map of the insert of pSP6K-mTRF23 and its sequence strategy. The insert of pSP6K-mTRF23 was isolated by *Bam*HI digestion and subcloned in pUC18. Clones that contained series of unidirectional deletions were produced by digestion with exonucleases III and VII from either end¹⁶. The nucleotide sequence of the entire cDNA was determined by the dideoxy method¹⁷. Top line, insert DNA; dotted and open rectangles, DNA sequences encoding the signal peptide and coding region, respectively. Horizontal arrows, directions and ranges of sequences read. Double arrow, 100 bp. *b*, *Bam*HI site. *c*, Nucleotide and predicted amino-acid sequences of TRF. Possible N-glycosylation sites are underlined. Numbers above and below sequences, amino-acid and nucleotide positions, respectively. *d*, Regions of identity between TRF and other lymphokines. Similar amino-acid sequences of known proteins were searched by computer program. Sequences (>14 residues) with 45% similarity (including related amino acids at the corresponding region) were chosen. Identical and related amino acids are shown by asterisks and dots, respectively. Sequences of murine GM-CSF²¹, murine interferon-γ (IFN-γ)²² and murine IL-3^{23,24} were taken from published details.

Xenopus oocytes and the translation products secreted into the incubation medium collected and assayed for proliferation and plaque-formation of BCL₁ cells. Translation products derived from *Sal*I-digested templates have both of these activities whereas those from *Sac*I-digested templates have neither, indicating that cDNAs encoding BCGFII and TRF contain *Sac*I sites in the structural sequences (Table 1).

The pSP6K cDNA library of 2.19 T-cell mRNA was divided into 18 pools, each consisting of 3,000 clones. We assayed each pool and found only one that was positive both in proliferation and in IgM plaque-forming activity of BCL₁ cells. After two more subpooling steps, we identified a positive pool of 60 clones and examined inserts of the individuals to select clones whose inserts were >1 kilobase (kb) and contained the *Sac*I site but not the *Sal*I site. Seven out of 60 clones (clones 16, 18, 23, 27, 42, 51 and 53) fulfilled these criteria. One of them, clone 23,

was positive for BCL₁ proliferation and IgM plaque-formation (Table 1) and was termed pSP6K-mTRF23.

The nucleotide sequence of the entire cDNA was determined by the combination of the unidirectional deletion method¹⁶ using exonucleases III and VII and the dideoxy method¹⁷. The cDNA insert is 1,533 base pairs (bp) long excluding the poly(A) tail (Fig. 2). There are two long open reading frames (ORFs) deduced from the cDNA sequence; the longer frame (133 residues) between bases 44 and 442 and the shorter (62) between bases 544 and 729. The TRF activity occurs in the translation products of pSP6K-mTRF23 linearized by *Acc*I (Table 1). Because the *Acc*I site is located in the middle of the shorter ORF at bases 624–629, the functional significance of this reading frame was excluded. The longer ORF encodes a polypeptide of 133 amino acids with three possible N-glycosylation sites. There are 21 hydrophobic amino acids at the N-terminus of this ORF

Table 3 BCGFII activity of recombinant TRF

Lymphokines	Additions (%)	c.p.m. per culture DXS -	c.p.m. per culture DXS +	Synergy indices
none		520	842	1.6
rTRF	5	2,715	6,044	2.2
rTRF	1.25	1,724	4,268	2.5
rTRF	0.6	1,290	2,813	2.2
rTRF	0.16	919	1,827	2.0
2.19 sup	3.0	7,425	11,214	1.5
2.19 sup	1.0	7,183	10,740	1.5
2.19 sup	0.3	6,424	8,626	1.3

CBA spleen B cells enriched by panning²⁰ were cultured in duplicates of $2 \times 10^6 \text{ ml}^{-1}$ in 0.2 ml RPMI1640 containing $50 \mu\text{M}$ 2-mercaptoethanol and 15% FCS with or without ($50 \mu\text{g ml}^{-1}$) dextran sulphate (DXS). Supernatants of oocytes injected with pSP6K-mTRF23 mRNA (rTRF) and culture supernatants of 2.19 T cell line (2.19 sup) were added at indicated concentrations. ^3H -thymidine was added for the last 24 h of a 72-h culture and the degree of proliferation was assessed by determining incorporation of the radiolabel. Synergy indices were calculated by dividing c.p.m. in DXS-stimulated cultures by c.p.m. in unstimulated culture.

(including two arginine residues). The putative secreted core polypeptide has a relative molecular mass (M_r) of 12,300, which is near that of TRF (M_r 18,000 after glycosylation¹²). The longer ORF contains three cysteine residues (15, 62 and 104) that may be involved in the oligomer formation. We therefore conclude that the longer ORF encodes the TRF molecule.

We searched homologies between the deduced amino-acid sequence of TRF and sequences of known proteins, including all the lymphokines so far identified. No proteins are very similar to TRF except for short segments of murine interleukin-3 (IL-3), murine granulocyte-macrophage colony stimulating factor (GM-CSF) and murine interferon- γ (Fig. 2c). We suggest that TRF is distantly related to these lymphokines.

We tested various biological activities of recombinant TRF (rTRF), the oocyte translation product of pSP6K-mTRF23. First, we tested it for the other activity of TRF, namely, stimulation of the secondary anti-DNP IgG response. Mice were immunized with DNP-keyhole limpet haemocyanin (KLH), and splenic B cells were purified 6–8 weeks later for the assay of the anti-DNP IgG plaque-forming cells. The assay revealed that rTRF stimulates DNP-primed B cells in the presence of antigen (DNP-OA) and augments anti-DNP IgG response (Table 2). The rTRF also induces IgM synthesis in *in vivo*-activated B-cell blasts (data not shown). These properties of rTRF agree with the criteria of TRF derived from B151 (refs 11–13).

During screening, we found that pools containing the stimulating activity of IgM plaque-forming cells of BCL₁ cells always have the stimulation activity of the BCL₁ proliferation. The translation product of the final clone, pSP6K-mTRF23, also shows these two activities (Table 1). Furthermore, rTRF augments ^3H -thymidine incorporation into B cells stimulated with dextran sulphate (Table 3). These results indicate that the TRF and BCGFII activities belong to a single molecule. The rTRF has no interleukin-1 (IL-1)-, IL-2-, IL-3- or BSF-1-like activity (data not shown), but it induces augmentation of ^3H -thymidine uptake in thymocytes in conjunction with suboptimal doses of Concanavalin A and IL-2. Although Sanderson *et al.*¹⁸ have recently suggested that BCGF-II is identical to eosinophil differentiation factor, we do not yet know whether TRF has this activity. Because of the diverse activities and targets of TRF, this lymphokine is classified as an interleukin. As this is the fifth interleukin whose structure has been determined, we propose that TRF or BCGF-II be called interleukin-5.

We thank Professor T. Oi for advice on the homology search and Ms K. Hirano for preparation of the manuscript. Supported by grants from the Ministry of Education, Science and Culture

of Japan, the Swedish National Science Research Council, the Swedish MRC and the Aristotale Onassis benefit foundation.

Received 23 June; accepted 20 August 1986.

- Dutton, R. W., Folkoff, R. & Hirst, J. A. *Prog. Immun.* **1**, 355–368 (1971).
- Swain, S. L. *et al. J. exp. Med.* **158**, 822–835 (1983).
- Howard, M., Nakanishi, K. & Paul, E. W. *Immun. Rev.* **78**, 185–210 (1984).
- Kishimoto, T. A. *Rev. Immun.* **3**, 133–157 (1985).
- Isakson, P. C., Puré, E., Vitetta, E. S. & Krammer, P. H. *J. exp. Med.* **155**, 734–748 (1982).
- Bergstedt-Lindqvist, S., Sideras, P., MacDonald, H. R. & Severinson, E. *Immun. Rev.* **78**, 25–50 (1984).
- Noma, Y. *et al. Nature* **319**, 640–646 (1986).
- Lee, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 2061–2065 (1986).
- Schimpl, A. & Wecker, E. *Nature new Biol.* **237**, 15–17 (1972).
- Takatsu, K., Tominaga, A. & Hamaoka, T. *J. Immun.* **124**, 2414–2422 (1980).
- Takatsu, K. *et al. J. Immun.* **125**, 2646–2653 (1980).
- Takatsu, K. *et al. J. Immun.* **134**, 382–389 (1985).
- Harada, N. *et al. J. Immun.* **134**, 3944–3951 (1985).
- Sideras, P., Bergstedt-Lindqvist, S., MacDonald, H. R. & Severinson, E. *Eur. J. Immun.* **15**, 586–593 (1985).
- Tominaga, A., Matsumoto, M., Takahashi, T., Harada, N. & Takatsu, K. *Microbiol. Immun.* (in the press).
- Yanisch-Perron, C., Vieira, J. & Messing, J. *Gene* **33**, 103–119 (1985).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5469 (1977).
- Sanderson, C., O'Garra, A., Warren, J. D. & Klaus, G. B. G. *Proc. natn. Acad. Sci. U.S.A.* **83**, 437–440 (1986).
- Gronowicz, E., Coutinho, A. & Melchers, F. *Eur. J. Immun.* **6**, 588–590 (1976).
- Mage, M. G., McHugh, L. L. & Rothstein, T. L. *J. immun. Meth.* **15**, 47–56 (1977).
- Gough, M. N. *et al. Nature* **309**, 763–767 (1984).
- Gray, W. P. & Goeddel, V. D. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5842–5846 (1983).
- Fung, C. M. *et al. Nature* **307**, 233–237 (1984).
- Yokota, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 1070–1074 (1984).

Complementary DNA for a novel human interleukin (BSF-2) that induces B lymphocytes to produce immunoglobulin

Toshio Hirano, Kiyoshi Yasukawa, Hisashi Harada*, Tetsuya Taga, Yasuo Watanabe, Tadashi Matsuda, Shin-ichiro Kashiwamura, Koichi Nakajima, Koichi Koyama, Akihiro Iwamatsu†, Susumu Tsunasawa†, Fumio Sakiyama†, Hiroshi Matsui‡, Yoshiyuki Takahara‡, Tadatsugu Taniguchi* & Tadimitsu Kishimoto

Division of Immunology and * Division of Molecular Biology, Institute for Molecular and Cellular Biology, † Division of Protein Chemistry, Institute for Protein Research, Osaka University, 1–3, Yamada-Oka, Suita, Osaka 565, Japan
‡ Central Research Laboratories, Ajinomoto Co. Inc., Totsuka-ku, Yokohama 244, Japan

When stimulated with antigen, B cells are influenced by T cells to proliferate and differentiate into antibody-forming cells. Since it was reported^{1,2} that soluble factors could replace certain functions of helper T cells in the antibody response, several different kinds of lymphokines and monokines have been reported in B-cell growth and differentiation^{3,4}. Among these, human B-cell differentiation factor (BCDF or BSF-2) has been shown to induce the final maturation of B cells into immunoglobulin-secreting cells^{5–8}. BSF-2 was purified to homogeneity⁹ and its partial NH₂-terminal amino-acid sequence was determined¹⁰. These studies indicated that BSF-2 is functionally and structurally unlike other known proteins. Here, we report the molecular cloning, structural analysis and functional expression of the cDNA encoding human BSF-2. The primary sequence of BSF-2 deduced from the cDNA reveals that BSF-2 is a novel interleukin consisting of 184 amino acids.

In a previous study¹¹, we established a human T-cell line (TCL-Na1), which constitutively produced large amounts of BSF-2, using human T-cell leukaemia virus-1 (HTLV-1). BSF-2, with a relative molecular mass (M_r) 21,000 (21K), was purified

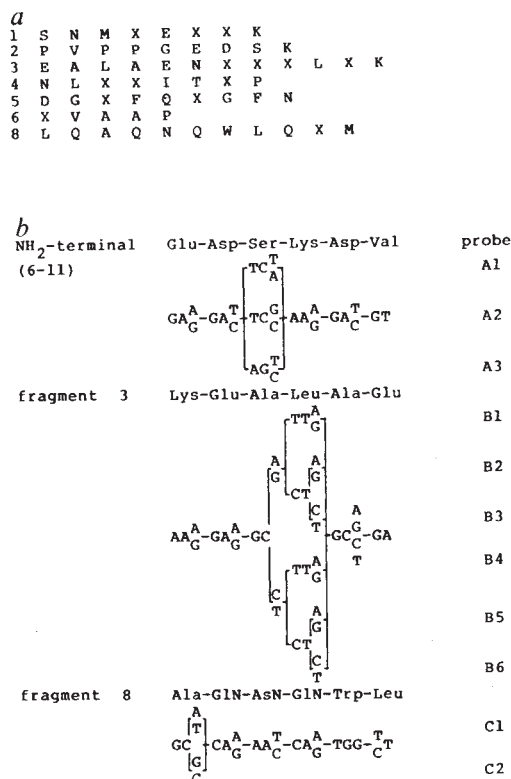


Fig. 1 *a*, Sequences of BSF-2 fragments; *b*, sequences of synthetic oligonucleotide probes.

Methods. *a*, BSF 2 (21K) was purified as reported previously^{9,10}. Approximately 1 nmol of purified BSF-2 (20 µg) was digested with 5 pmol of lysylendopeptidase in 5 mM Tris-HCl buffer (pH 9.5) at 37 °C for 6 h. The fragments were separated by high performance liquid chromatography (HPLC) using reverse phase column, µBondpack (C₁₈ 300 × 2.1 mm, Waters), with a linear gradient of 0–60% acetonitrile containing 0.1% trifluoroacetic acid. The partial amino-acid sequence of each fragment was determined using a gas-phase protein sequencer (Applied Biosystems model 470A) essentially as described¹⁷. The obtained phenylthiohydantoin-amino acid residues were analysed by HPLC using ODS-120T column¹⁸ (C₁₈ 250 × 4.6 mm, Toyo Soda). *b*, Mixtures of 32 (A1–A3, C1 and C2) or 64 (B1–B6) possible oligonucleotides (17 bases) corresponding to the NH₂-terminal (A1–A3), fragment 3 (B1–B6) and fragment 8 (C1, C2) were made by an automated DNA synthesizer (Applied Biosystem Inc.). The expression vector pQ was constructed by a series of manipulations of pCEIL-21¹⁹. Essentially, it contains two copies of a simian virus 40 (SV40) DNA segment containing the early promoter and *ori* sequences in opposite orientations. This plasmid is unstable unless a DNA (in this case cDNA) segment is inserted between the two SV40 sequences, because it becomes cruciform in *Escherichia coli*²⁰. Once the cDNA is inserted it can be expressed in monkey COS cells independent of orientation. Efficiency of expression was previously confirmed with IL-2 and IFN cDNA (G. Yamada and T. Taniguchi, unpublished data). Poly (A)⁺ RNA was prepared from TCL-Na1 cells and cDNA was made using the Amersham cDNA synthesis system (Amersham UK) which is essentially based on the method described by Gubler and Hoffman²¹. pQ vector was digested with *Pst*I to remove the insert IL-2 cDNA. One part of the cDNA was annealed with pQ vector at the *Pst*I site by the standard G–C tailing method¹⁹ and the resulting recombinant plasmids were transfected into *E. coli* MC1061. The other part of the cDNA was methylated with *Eco*RI methylase (New England Biolabs), ligated with *Eco*RI linker (Takara Shuzo Co.), digested with *Eco*RI and applied to Bio-Gel A-50 m (Bio-Rad). cDNA fragments longer than 500 base pairs were ligated with *Eco*RI-digested λgt10 vector (Stratagene) and packaged *in vitro* using Gigapack (Stratagene)²². Both cDNA libraries (30 × 10⁴ clones) were screened with the ³²P-labelled 17-base-long oligonucleotides by the *in situ* procedure under the conditions described²³.

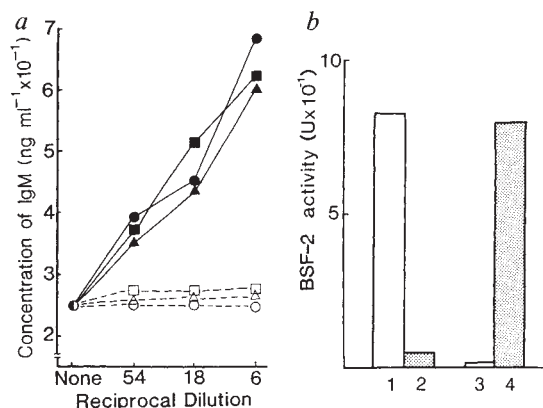
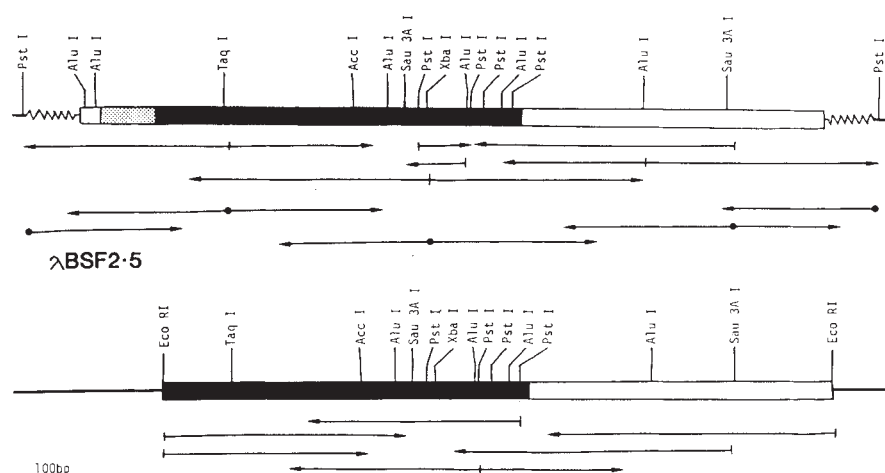


Fig. 2 *a*, Expression of BSF-2 cDNA in COS7 cells; *b*, affinity chromatography of COS7 cell-derived BSF-2.

Methods. Monkey COS7 cells²⁴ (5×10^5) were seeded onto 60 mm plates one day before transfection. Transfections were carried out with 2 µg of plasmid DNA in 1 ml 25 mM Tris-HCl (pH 7.4), 137 mM NaCl, 5 mM KCl, 0.6 mM Na₂HPO₄, 0.5 mM MgCl₂, 0.7 mM CaCl₂ and 500 µg of DEAE-Dextran (Pharmacia). After 1 h incubation at 37 °C, the solution was replaced with Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum (FCS) and 150 µM chloroquine, incubated at 37 °C for 3 h and replaced with fresh DMEM containing 10% FCS. Seventy-two hours later the culture supernatants were collected and assayed for BSF-2 activity. *a*, Ten thousand cells of the Epstein-Barr virus transformed B cell line SKW6-C14 were cultured in different concentrations of either the culture supernatants of COS7 cells transfected with pBSF2.38 DNA (●, ■, ▲) or mock-transfected COS7 cells (○, □, △). Three separate experiments are shown. After 3 days the levels of IgM in the culture supernatants were determined by enzyme-linked immunosorbent assay (ELISA) as described⁹. *b*, The culture supernatants of COS7 cells transfected with pBSF2.38 DNA (containing 90 units of BSF-2) were applied to specific anti-peptide immunoaffinity gel (2, 4) or to control anti-OVA immunoaffinity gel (1, 3). After washing the gels, the bound protein was eluted with 3M KSCN. BSF-2 activities in the effluent (1, 2) and eluate (3, 4) were determined using SKW6-C14 cells. One unit per ml of BSF-2 induces 50% of the maximum IgM-production in 1×10^4 SKW6-C14 cells⁹. The specific immunoaffinity gel was made using affinity-purified rabbit antibody (anti-peptide) against the synthetic NH₂-terminal BSF-2 peptide¹⁰.

to homogeneity from the culture supernatant of TCL-Na1 cells and digested with lysylendopeptidase to give 9 fragments. The partial amino acid sequences of 7 fragments were determined (Fig. 1a). On the basis of the partial amino acid sequence at the NH₂-terminal of BSF-2 reported previously¹⁰, it was shown that fragment 2 consisted of 9 amino acids at the NH₂-terminal and fragment 6 consisted of residues starting with the tenth amino acid from the NH₂-terminal. The cDNA libraries made from poly(A)⁺ RNA of TCL-Na1 cells were probed with 11 groups of synthetic oligonucleotide mixtures (17 bases each) corresponding to fragments 3, 8 and the NH₂-terminal as indicated in Fig. 1b. One cDNA clone, designated pBSF2.38, that specifically hybridized with A2, B2 and C1 probes but no others, was isolated from the cDNA library constructed in pQ vector. The libraries were screened again by pBSF2.38 insert cDNA as a probe and ten positive clones were identified. Four of these were shown to hybridize with the A2 probe, which corresponds to the NH₂-terminal of BSF-2. All five cloned plasmids contained cDNA inserts consisting of about 1,000–1,100 base pairs. To examine whether pBSF2.38 contained the entire coding region of BSF-2, pBSF2.38 DNA was transfected into COS7 cells. The culture supernatant of COS7 cells transfected with pBSF2.38 DNA induced IgM production in SKW6-C14 cells and the control culture supernatant of mock-transfected COS7 cells did not (Fig. 2a). The BSF-2 activity generated by COS7 cells transfected with pBSF2.38 DNA could be adsorbed onto and

a pBSF2.38



subsequently eluted from an immunoaffinity gel conjugated with polyclonal antibody against the synthetic NH₂-terminal BSF-2 peptide, but did not bind to the control antiovalbumin (OVA) immunoaffinity gel (Fig. 2b). Thus pBSF2.38 contains the entire coding region of BSF-2.

The complete nucleotide sequence of the insert cDNA of pBSF2.38 was determined by the dideoxy method¹² and by chemical cleavage¹³. The restriction endonuclease cleavage map of the cDNA insert and the sequencing strategy are shown in Fig. 3a. The insert cDNA of pBSF2.38 contains a single large open reading frame (Fig. 3b). In this frame, the initiator ATG is followed by 211 codons before the termination triplet TAG. We confirmed the nucleotide sequence of BSF-2 cDNA using one of the clones (λ BSF2.5) identified by probing the cDNA library constructed in λ gt 10 vector with pBSF2.38 insert DNA (Fig. 3a). Although λ BSF2.5 insert cDNA was found to lack 126 nucleotides at the 5' terminal region compared with pBSF2.38 insert cDNA, the nucleotide sequence of the other region was in complete agreement with that of pBSF2.38 cDNA insert (data not shown). Because we know from protein sequence data that the sequence of the NH₂-terminal region of mature BSF-2 is Pro-Val-Pro-Pro . . . , we conclude that BSF-2 is made as a precursor consisting of 212 amino acids and processed into a mature form by a cleavage between the Ala (-1) and Pro (+1) residues. Hydrophobic amino acids are abundant in residues from -28 to -1 and this region seems to be a typical signal peptide that is required for BSF-2 secretion. Although we do not know at present whether the carboxy-terminal processing occurs in BSF-2, we conclude that mature BSF-2 probably consists of 184 amino acids. As underlined in Fig. 3b, the amino-acid sequence of all the fragments that originated from purified BSF-2 was in complete agreement with the amino-acid sequence deduced from the cDNA. The calculated molecular mass of the BSF-2 polypeptide was 20,781.5. This is comparable with the apparent molecular mass of purified BSF-2 (21K) as determined by SDS-polyacrylamide gel electrophoresis (PAGE)^{9,10}. Two potential N-glycosylation sites (Asn-X-Ser and Asn-X-Thr, at positions 45-47 and 144-146, respectively) were identified, but we have not yet determined whether BSF-2 is a glycoprotein.

The protein sequence of BSF-2 was compared with all proteins in the National Biomedical Research Foundation database and Genetic Sequence Data Bank (Bolt Beranek and Newman, Inc., Los Alamos National Laboratory); no strong similarity with any protein was revealed. We also searched our personal database, which contains human IL-1 α , IL-1 β , IL-2, γ -interferon (γ -IFN), colony stimulating factor-1 (CSF-1), granulocyte macrophage (GM)-CSF, granulocyte (G)-CSF and B-cell stimulatory factor-1 (BSF-1), and murine IL-1, IL-2, γ -IFN, GM-CSF, IL-3 and

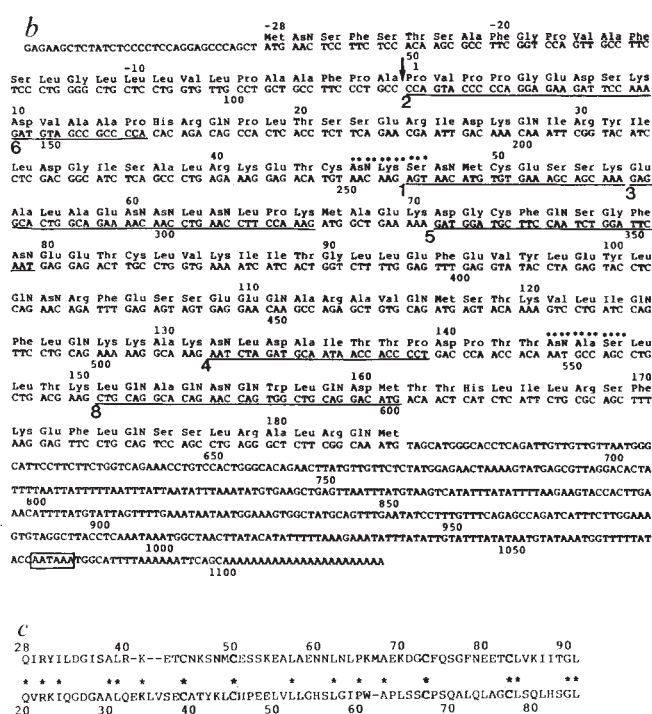


Fig. 3 *a*, Restriction endonuclease cleavage map and sequence strategy for the cDNA insert of pBSF2.38 and λ BSF2.5. *b*, Nucleotide sequence and deduced amino-acid sequence of the plasmid pBSF2.38 cDNA insert. *c*, Homology of human BSF2 and G-CSF. *a*, Arrows with vertical marks indicate the direction and extent of sequencing by the dideoxy chain termination method. Arrows with dots, regions of DNA sequenced by chemical cleavage. Scale, 100 bp. *b*, Underlined amino acids, the region matching the sequence of the lysylenopeptidase-digested peptides 1-6 and 8 (see Fig. 1*a*). All derived peptide sequences are preceded by a lysine, consistent with lysylenopeptidase digestion. Asterisks, potential *N*-glycosylation sites (Asn-X-Ser/Thr); box, presumed poly(A) addition signal sequence. Numbers above and below the sequence show positions of amino acids and nucleotides, respectively. Amino acids are numbered starting at Pro (1) of the mature BSF-2 protein sequence. *c*, Homology search was performed using the search program of Wilbur and Lipmann²⁵. The parameters used in the database search were κ -tuple size = 2, window size = 20, weight for a deletion = 8 and value to be added to each element of the similarity matrix = 0. Under these conditions, only human G-CSF showed significant similarity to BSF-2 (homology score = -129). Asterisks, homologous amino acids; shaded boxes, cysteine. Sequence of G-CSF was taken from Nagata *et al.*¹⁴.

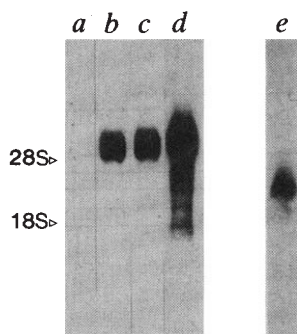


Fig. 4 Analysis of BSF-2 mRNA. Poly(A)⁺ RNA was prepared from various kinds of cells. RNA (10 µg) was denatured and subjected to blotting analysis²⁶. *a*, Poly(A)⁺ RNA from non-stimulated tonsillar mononuclear cells; *b*, poly(A)⁺ RNA from tonsillar mononuclear cells stimulated with phytohaemagglutinin (0.1%) and tetradecanoyl phorbol 13-acetate (5 ng ml⁻¹) for 40 h; *c*, poly(A)⁺ RNA from TCL-Na1 cells; *d*, poly(A)⁺ RNA from T 24 bladder cell carcinoma; *e*, total RNA from cardiac myxoma cells¹⁰.

BSF-1 for similarities. Only human G-CSF¹⁴ showed significant similarity to BSF-2. Amino acids 28–91 of BSF-2 and amino acids 20–85 of G-CSF match at 17 residues (25.7% of all residues) (Fig. 3c). Note that cysteine residues at positions 44, 50, 73 and 83 of BSF-2 match those at positions 39, 45, 67 and 77 of G-CSF exactly. This suggests that BSF-2 is distantly related to G-CSF. If so, intramolecular disulphide bonds would be important in the structure of these proteins.

In Northern blot analysis (Fig. 4) pBSF2.38 DNA hybridized with a single species of messenger RNA ~1,300 nucleotides long which was extracted from activated lymphocytes and TCL-Na1 cells but not from unstimulated control lymphocytes, showing that, like other T-cell-derived lymphokines, BSF-2 mRNA is induced upon mitogenic stimulation.

We have previously reported¹⁰ that cardiac myxoma cells produce high levels of a molecule with BSF-2 activity which was specifically absorbed by the anti-peptide antibody. It has also been shown that bladder cell carcinoma, T24 cells produce a molecule with BSF-2 activity¹⁵. To show that these cells produce BSF-2 itself, we performed Northern blot analysis. As shown in Fig. 4 (*d, e*), BSF-2 mRNA was transcribed in T24 cells and myxoma cells at a much higher level than activated lymphocytes, although the mRNA in myxoma cells was larger than that of lymphocytes. These data indicate that under certain conditions, the BSF-2 gene can be expressed in non-lymphoid tissues. Because patients with cardiac myxoma frequently show several kinds of autoimmune disease-like symptoms¹⁶, it was suggested¹⁰ that deregulated production of BSF-2 might be involved in autoimmune pathogenesis. The fact that BSF-2 gene expression in myxoma cells is high further strengthens this argument.

This work was supported in part by grants from the Ministry of Education, Science and Culture. We thank Dr Edward Barsumian for critical review of the manuscript and Ms K. Kubota, J. Mori and M. Kawata for their secretarial assistance.

Received 4 August; accepted 26 September 1986.

1. Dutton, R. W. *et al. Prog. Immun.* **1**, 355–368 (1971).
2. Schimpl, A. & Wecker, E. *Nature* **237**, 15–17 (1972).
3. Howard, M. & Paul, W. E. *A. Rev. Immun.* **1**, 307–333 (1983).
4. Kishimoto, T. *A. Rev. Immun.* **3**, 135–159 (1985).
5. Okada, M. *et al. J. exp. Med.* **157**, 583–590 (1983).
6. Butler, J. L., Falkoff, R. J. M. & Fauci, A. S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2475–2478 (1984).
7. Hirano, T., Teranishi, T., Lin, B. H. & Onoue, K. *J. Immun.* **133**, 798–802 (1984).
8. Kikutani, H. *et al. J. Immun.* **134**, 990–995 (1985).
9. Hirano, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5490–5494 (1985).
10. Hirano, T. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
11. Shimizu, K. *et al. J. Immun.* **133**, 1728–1733 (1985).
12. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
13. Maxam, A. W. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
14. Nagata, S. *et al. Nature* **319**, 415–418 (1986).

15. Rawle, F. C. *et al. Eur. J. Immun.* **16**, 1017–1019 (1986).
16. Sutton, M. G. S. J., Mercier, L., Giuliani, E. R. & Lie, J. T. *Mayo Clin. Proc.* **55**, 371–376 (1980).
17. Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. *J. Biol. Chem.* **256**, 7990–7997 (1981).
18. Tsunasawa, S., Kondo, J. & Sakiyama, F. *J. Biochem. Tokyo* **97**, 701–704 (1985).
19. Taniguchi, T. *et al. Nature* **302**, 305–310 (1983).
20. Mizuuchi, K., Mizuuchi, M. & Gellert, M. *J. molec. Biol.* **156**, 229–243 (1982).
21. Gubler, U. & Hoffman, B. J. *Gene* **25**, 263–269 (1983).
22. Huynh, T. V., Young, R. A. & Davis, R. W. *DNA Cloning*. Vol. 1, 49–78 (IRL press, Oxford, 1985).
23. Wood, W. I., Gitshier, J., Lasky, L. A. & Lawn, R. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1585–1588 (1985).
24. Gluzman, Y. *Cell* **23**, 175–182 (1981).
25. Wilbur, W. J. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 726–730 (1983).
26. Lehrach, H., Diamond, D., Wozey, J. M. & Boedtker, H. *Biochemistry* **16**, 4743–4751 (1977).

Intracellularly injected tetanus toxin inhibits exocytosis in bovine adrenal chromaffin cells

Reinhold Penner*, Erwin Neher† & Florian Dreyer*

* Rudolf-Buchheim-Institut für Pharmakologie, Frankfurter Str. 107, D-6300 Giessen, FRG

† Max-Planck-Institut für Biophysikalische Chemie, Am Fassberg, D-3400 Göttingen, FRG

The clostridial neurotoxins tetanus and botulinum toxin type A are known to block transmitter release from nerve terminals^{1–3}, probably by interfering with some essential process controlling exocytosis^{3,4} after the entry of Ca²⁺ ions. Although exocytosis occurs in many secretory cells, these toxins show a high specificity for neurones and the secretory response of cultured bovine adrenal medullary cells is not inhibited by exposure to medium containing tetanus or botulinum toxin type A (although it is by botulinum toxin type D)⁴. Here we report that when tetanus toxin and botulinum neurotoxin type A are injected intracellularly into chromaffin cells they strongly inhibit secretion, as revealed by the measurement of cell capacitance⁵. These results indicate that these toxins are normally ineffective in chromaffin cells because they are not bound and internalized, so do not reach their site of action. Furthermore, we have localized the secretion-blocking effects of the toxin to a fragment comprising the light chain covalently linked to part of the heavy chain, suggesting that this part of the molecule contains the active site.

We examined cultured bovine adrenal chromaffin cells, isolated as described⁶, using the whole-cell configuration of the patch-clamp technique⁷. This allows one to monitor exocytosis using membrane capacitance as a measure of the cell surface area. Capacitance and surface area change considerably as chromaffin granules fuse with the membrane or endocytotic retrieval of the membrane occurs⁵. Secretion was stimulated by dialysing the cells through a patch-pipette with a solution buffered to a free Ca²⁺ concentration of 1 µM. The total membrane capacitance of control cells (typically 4–8 pF) started to increase immediately upon establishment of the whole-cell configuration, eventually leading to a 2–3-fold capacitance increase within 15–25 min (Fig. 1a).

When cells were preloaded with either tetanus toxin or botulinum neurotoxin type A the secretory response to the same challenge (dialysis of cells with Ca²⁺-containing solution) was effectively removed. In fact, cells pretreated with tetanus toxin consistently showed a slight decrease in total membrane capacitance (Fig. 1c), suggesting that in these cells there was more endocytotic than exocytotic activity. This 'reversal effect' was observed only once in cells pretreated with botulinum neurotoxin type A (*n* = 10) which may reflect genuine differences in the ability of the two toxins to inhibit secretion or, as has been suggested^{8,9}, different sites of action. The typical response to increased intracellular Ca²⁺ levels in cells preloaded with botulinum neurotoxin type A was a slight initial increase in cell

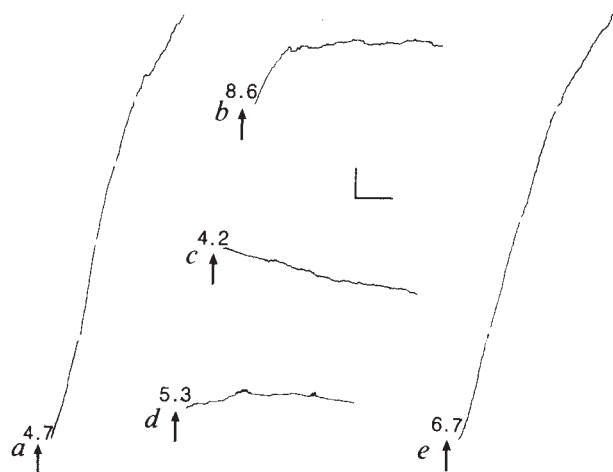


Fig. 1 Effects of clostridial toxins on the secretory responses (measured as capacitance changes) of bovine chromaffin cells stimulated by elevation of the intracellular Ca^{2+} concentration. *a*, Typical time course of capacitance changes recorded from a stimulated control cell. *b*, Typical inhibitory effect of botulinum A toxin on the response of a stimulated cell. *c*, Inhibition of exocytosis by tetanus toxin. *d*, Inhibition of exocytosis in cells preloaded with fragment B at $25 \mu\text{g ml}^{-1}$ in the pipette. *e*, Capacitance increase recorded from a cell preloaded with fragment C at $25 \mu\text{g ml}^{-1}$ in the pipette. Capacitance measurements were performed according to the method described by Neher and Marty⁵. Briefly, adrenal medullary chromaffin cells isolated as described⁶ and kept in short-term culture (1–5 days) were voltage clamped at -70 mV and subjected to a sinusoidal voltage command of 18 mV r.m.s. at 795 Hz . The resulting current was fed into a lock-in amplifier which delivered the capacitive component of the signal when tuned to the desired phase angle with respect to the reference signal. For the large capacitance changes of the control and fragment C traces, these had to be reconstructed from several segments (interruptions) since the lock-in measurement requires recompensation after large deflections⁵. In all traces establishment of the whole-cell recording configuration is indicated by arrows. Numbers given at the start of each trace correspond to initial cell capacitance values (in pF) as determined by the dial setting of the EPC-7 amplifier (List Electronic, Darmstadt). Average capacitance values were $5.9 \pm 0.3 \text{ pF}$ (mean $\pm$ s.e.m., $n=25$), series resistance values, as read from the EPC-7 settings, were $3.9 \pm 0.5 \text{ M}\Omega$. *a*, To stimulate secretion a pipette filling solution of the following composition was used (mM): K glutamate 135, NaCl 20, MgCl_2 4, ATP 0.5, EGTA 10, CaCl_2 9 (free Ca^{2+} $1 \mu\text{M}$), HEPES 10 (pH 7.2). The bath solution contained (mM): NaCl 140, KCl 2.8, MgCl_2 2, CaCl_2 2, HEPES 10 (pH 7.2). The cell underwent the pretreatments described in *b* except that toxin was omitted from the pipette when 'loading' the cell. *b*, The cell was loaded with the toxin in the whole-cell configuration for 3 min and incubated for 1 h at 37°C . Botulinum toxin ($\text{LD}_{50}=8 \text{ ng kg}^{-1}$, mouse s.c.) was used at a concentration of $50 \mu\text{g ml}^{-1}$ in a pipette filling solution containing (mM): K glutamate 135, NaCl 20, MgCl_2 4, EGTA 1, HEPES 10 (pH 7.2). Both the procedure of loading and the actual test for the secretory response were carried out at 30°C . *c*, Cells were pretreated, incubated and challenged as in *b* except when loading the cells $50 \mu\text{g ml}^{-1}$ tetanus toxin ($\text{LD}_{50}=3 \text{ ng kg}^{-1}$, mouse s.c.) was present in the electrode filling solution. Vertical scale bar, 0.2 pF ; horizontal scale bar, 1 min.

capacitance, followed by a phase in which no major alteration was seen (Fig. 1*b*). Cells preloaded using pipettes containing no toxin or denatured (boiled) toxins showed a normal secretory response. The rates of release resulting from an elevation of the intracellular Ca^{2+} of control and toxin-treated cells are shown in Fig. 2.

Both secretion and tetanus toxin effects are temperature dependent. Lowering the temperature slows, and can even prevent, the inhibitory effects of tetanus toxin at nerve terminals¹⁰. We therefore carried out experiments at different temperatures. Whereas secretion at 30°C proceeded at an average rate of 19.7% per min (see Fig. 2), at 20°C it was reduced to $7.5 \pm 1.4\%$ per min (mean $\pm$ s.e.m., $n=6$). At 20°C the release rate of cells pretreated with tetanus toxin ($5.7 \pm 1.5\%$ per min, mean $\pm$ s.e.m., $n=11$) was not significantly different from that of control cells. This finding clearly indicates that the temperature-dependent step in the action of tetanus toxin is subsequent to binding and internalization. Combined with the data obtained using boiled toxins, these results strongly support the conclusion that the

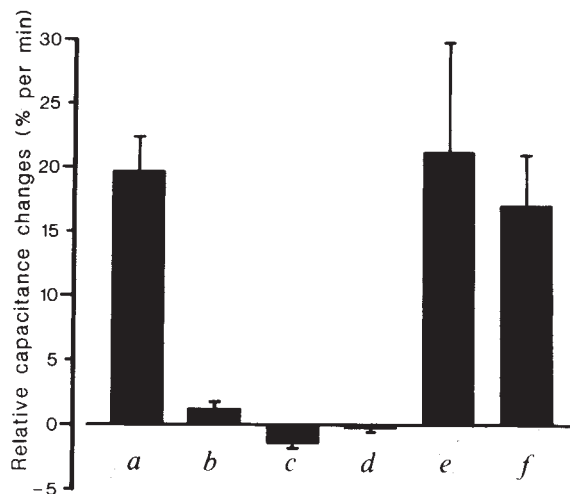


Fig. 2 Effects of clostridial toxins on secretion from bovine adrenal medullary cells. *a*, Control (mean $\pm$ s.e.m. of 14 cells); *b*, botulinum toxin A (10 cells); *c*, tetanus toxin (5 cells); *d*, fragment B (7 cells); *e*, fragment C (4 cells); *f*, denatured tetanus toxin (3 cells). The measure of release chosen was the relative capacitance increase (% of the initial value) per unit time (min). Care was taken to assure relatively uniform cell sizes so that differences in initial capacitance values for each set of experiments were small. To determine the release rates capacitance values were measured 2–4 min after transition from cell-attached to whole-cell configuration which usually corresponded to the time in which the capacitance increase was linear (see Fig. 1*a* and *e*).

inhibitory actions of tetanus and botulinum A toxin on the release mechanism are highly specific and not simply due to the intrusion of a protein of high relative molecular mass (M_r) into the cell.

We used the same system to test for intracellular effects of fragments B and C of tetanus toxin. Fragment C is a part of the toxin's heavy chain of M_r 50,000 which is thought to mediate binding of the toxin to cell membranes^{11,12}. Fragment B consists of the light chain covalently linked to the remaining part of the heavy chain (M_r 100,000). These fragments are obtained by enzymatic treatment with papain, and are non-toxic at doses of 10 mg kg^{-1} when injected into mice (the LD_{50} for tetanus toxin is 3 ng kg^{-1} , mouse subcutaneous (s.c.))¹³. Fragment C did not inhibit the responses of chromaffin cells to increased intracellular Ca^{2+} levels (Figs 1*e* and 2), but fragment B effectively suppressed the secretory response (see Figs 1*d* and 2), suggesting that this fragment bears the active site for blocking exocytosis. We have thus shown that a fragment of tetanus toxin that is otherwise non-toxic is able to block secretion.

It is not known whether fragmentation of the toxin into B and C occurs fortuitously or whether it is integral to the inhibitory actions of tetanus toxin. It appears that some fragmentation occurs *in vivo*, but most of the toxin extracted from the spinal cord (the destination of toxin after uptake into motor nerve terminals and retrograde transport) is indistinguishable from native toxin¹⁴. This suggests that the proposed uptake mechanism of clostridial neurotoxins² (drawn by analogy with diphtheria toxin and Semliki Forest virus) by receptor-mediated endocytosis followed by expulsion of the toxin's light chain through a 'channel' formed by the toxin's heavy chain into the cytoplasm may not apply. From our method of toxin administration it is clear that, even if this described uptake mechanism is used, any specific structural changes occurring during uptake are not absolutely required for the actions of the toxins. If changes that are required for activity occur, our results suggest that they are more likely to happen within the cytoplasm.

These results support the hypothesis that tetanus and botulinum A toxin act beyond the rise of intracellular Ca^{2+}

concentration, interfering with an as yet unknown step late in the chain of events leading to exocytosis. Our data show that this crucial step is not restricted to the release process from nerve terminals, but may be quite generally associated with vesicular release. Clostridial neurotoxins may serve as valuable tools in studying late steps in exocytosis. The whole-cell recording technique combined with capacitance measurement is useful because it allows the toxins to reach their site of action directly in cells which lack the prerequisite membrane binding and internalization features and allows tests of fragments that are not normally internalized. Experiments using smaller fragments may allow the characterization of a 'minimal toxin' small enough to be introduced into high-voltage¹⁵ or drug-permeabilized^{16,17} cells.

We thank Mr U. Weller and Professor E. Habermann for supplying purified fragments B and C. Tetanus toxin was from the Behringwerke (Marburg, FRG), botulinum toxin type A was a gift from Dr Schantz (Food Research Institute, Madison). We also thank Dr S. DeRiemer for valuable comments on the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 47).

Received 7 July; accepted 1 September 1986.

1. Mellanby, J. J. *Neurosci.* **11**, 29-34 (1984).
2. Simpson, L. L. A. *Rev. Pharmac. Tox.* **26**, 427-453 (1986).
3. Habermann, E. & Dreyer, F. *Curr. Top. Microbiol. Immun.* **129**, 93-179 (1986).
4. Knight, D. E., Tonge, D. A. & Baker, P. F. *Nature* **317**, 719-721 (1985).
5. Neher, E. & Marty, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6712-6716 (1982).
6. Fenwick, E. M., Fajdiga, P. B., Howe, N. B. S. & Livett, B. G. *J. Cell Biol.* **76**, 12-30 (1978).
7. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85-100 (1981).
8. Dreyer, F. & Schmitt, A. *Neurosci. Lett.* **26**, 307-311 (1981).
9. Dreyer, F. *et al. J. Physiol. (Paris)* **79**, 252-258 (1984).
10. Schmitt, A., Dreyer, F. & John, Chr. *Naunyn-Schmiedeberg's Archs Pharmac.* **317**, 326-330 (1981).
11. Morris, N. P. *et al. J. biol. Chem.* **255**, 6071-6076 (1980).
12. Goldberg, R. L., Costa, T., Habig, W. H., Kohn, L. D. & Hardegger, M. C. *Molec. Pharmac.* **20**, 565-570 (1981).
13. Weller, U., Taylor, C. F. & Habermann, E. *Toxicon* (in the press).
14. Habermann, E., Wellhöner, H. H. & Räker, K. O. *Naunyn-Schmiedeberg's Archs Pharmac.* **299**, 187-196 (1977).
15. Baker, P. F. & Knight, D. E. *Nature* **276**, 620-622 (1978).
16. Dunn, L. A. & Holz, R. W. *J. biol. Chem.* **258**, 4989-4993 (1983).
17. Wilson, S. P. & Kirshner, N. *J. biol. Chem.* **258**, 4994-5000 (1983).

The essential light chains constitute part of the active site of smooth muscle myosin

Yoh Okamoto*, Takamitsu Sekine*, Jean Grammer† & Ralph G. Yount†

* Department of Biochemistry, Juntendo University School of Medicine, Tokyo, Japan

† Biochemistry/Biophysics Program, Institute of Biological Chemistry and Department of Chemistry, Washington State University, Pullman, Washington 99164-4630, USA

Myosin, a major contractile protein, characteristically possesses a long coiled-coil α -helical tail and two heads. Each head contains both an actin binding site and an ATPase site and is formed from the NH₂-terminal half of one of the two heavy chains (relative molecular mass, M_r , 200,000) and a pair of light chains; the so-called regulatory and essential light chains of approximately M_r 20,000 each. Recently we have identified¹ Trp 130 of the myosin heavy chain from rabbit skeletal muscle as an active-site amino-acid residue after labelling with a new photoaffinity analogue of ADP, *N*-(4-azido-2-nitrophenyl)-2-aminoethyl diphosphate (NANDP)². Nonspecific labelling was eliminated by first trapping NANDP at the active site with thiol crosslinking agents³. Exclusive labelling of the heavy chains with no labelling of the light chains agreed with previous findings^{4,5} that the heavy chains alone

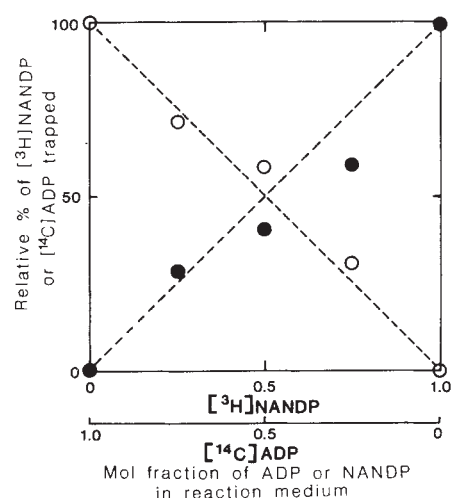


Fig. 1 The relative loss of Ca²⁺ and K⁺-EDTA ATPase activities of gizzard myosin as a function of vanadate-trapped [³H]NANDP. Gizzard myosin (4 mg ml⁻¹) in 0.5 M KCl, 2.0 mM MgCl₂, 20 mM Tris-Cl, pH 8.0, was incubated with 0.7 mM vanadate and varying concentrations of [³H]NANDP^{1,2} for 2 h at 0 °C. Aliquots (0.9 ml) were passed over Sephadex G-25 (PD-10) columns to remove vanadate and untrapped [³H]NANDP. The peak containing myosin-[³H]NANDP-vanadate was assayed for protein (biuret method) and for Ca²⁺ (○) and K⁺-EDTA (●) ATPase activities¹⁸. The [³H]NANDP content was determined by scintillation counting. Actin-free gizzard myosin was prepared from fresh chicken gizzards essentially by the procedure of Ebashi³⁰ and was in the dephosphorylated form. In addition, all experiments were performed at high ionic strength where all the myosin molecules are in the 6S unfolded state¹⁶. Vanadate solutions were prepared and analysed according to Goodno³¹.

contain the actin-activated Mg-ATPase activity of rabbit skeletal myosin. Here we report similar photolabelling experiments with smooth muscle myosin (chicken gizzard) in which ³H-NANDP is trapped at the active site with vanadate⁶ and which show that both the heavy chains and the essential light chains are labelled. The results indicate that both chains contribute to the ATP binding site and represent the first direct evidence for participation of the essential light chains in the active site of any type of myosin.

It is of primary importance in photoaffinity labelling studies to eliminate nonspecific labelling. The discoveries that nucleotides and nucleotide analogues could be trapped in the active site of myosin either by crosslinking two reactive thiols, SH₁ and SH₂ (refs 3, 7) or by forming a stable nucleotide complex with vanadate⁶ have given highly effective ways to achieve specific photolabelling. Both approaches lead to an ~10⁴ decrease in the off-rate of bound nucleotide and give complexes with lifetimes of several days at 0 °C. It is likely that both types of trapping yield quite similar active-site structures since we have found that binding of NANDP can be stabilized at the active site of cardiac myosin by either method (Y.O. and R.G.Y., manuscript in preparation). Photolysis in each case resulted in 50-60% covalent photolabelling of a heavy-chain peptide which was almost identical in sequence to that obtained from rabbit skeletal myosin¹. Similar studies with other ATP photoaffinity analogues all indicate that heavy chains of skeletal myosin and not light chains are modified^{7,8}. Earlier studies by Szilagyi *et al.*⁹ with an analogue containing a nitroarylazide group on the ribose ring of ATP also photolabelled predominantly the heavy chains although a small amount of labelling on the light chains could not be ruled out. These studies did not, however, utilize either of the two trapping methods mentioned above.

It was of interest to extend our studies to smooth muscle myosin because these molecules, unlike myosins from skeletal and cardiac muscles, exhibit myosin-linked regulation of contraction^{10,11}. In addition, gizzard myosin is known to undergo

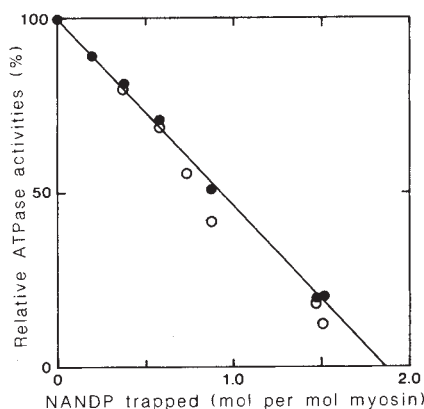


Fig. 2 Competitive trapping of [^{14}C]ADP (○) or [^3H]NANP (●) on gizzard myosin by use of the vanadate ion. Myosin (4 mg ml^{-1}) in 0.5 M KCl , 2.0 mM MgCl_2 , 20 mM Tris-Cl , $\text{pH } 8.0$, was incubated with 1 mM vanadate and varying ratios of [^{14}C]ADP (Amersham) to [^3H]NANP (total concentration of ADP plus NANP = 0.1 mM) for 2 h at 0°C . The myosin-[^{14}C]ADP (or [^3H]NANP)-vanadate complexes were purified and analysed as described in Fig. 1. The maximal amount of trapping under these conditions was $1.34\text{ mol } [^3\text{H}] \text{NANP}$ per mol myosin and $1.42\text{ mol } [^{14}\text{C}] \text{ADP}$ per mol myosin.

large changes in conformation as a function of phosphorylation of the regulatory light chain (LC_{20}) (refs 12–14), the binding of Mg-ATP^{15} , or changes in ionic strength^{16,17}. In particular, it seemed possible that specific photoprobes could reveal changes in the composition of the active site as a function of these induced conformational changes. Moreover, earlier studies using modified gizzard myosin¹⁸ had indicated a close connection between ATP cleavage and changes in a fluorescently-labelled essential light chain, suggesting that essential light chains (LC_{17}) were closely linked to the active site. We trapped [^3H]NANP on gizzard myosin with a vanadate ion in a manner similar to that described by Goodno for ADP on skeletal myosin⁶. In this work, as shown in Fig. 1, vanadate stabilized the binding of [^3H]NANP to the active site of gizzard myosin as judged by the parallel loss of both the Ca^{2+} - and K^+ -EDTA ATPase activities and by the nearly stoichiometric trapping of [^3H]NANP to each head.

Competitive vanadate trapping experiments were done with [^{14}C]ADP and [^3H]NANP to verify further that NANP was binding at the active site. Figure 2 shows that as the mol fraction of [^{14}C]ADP was increased relative to the mol fraction of [^3H]NANP, more [^{14}C]ADP was trapped but that the total amount of diphosphate compounds trapped remained constant. These results indicate that [^3H]NANP and [^{14}C]ADP bind with comparable affinities to a single type of site and that they compete effectively with each other for stable binding to gizzard myosin.

The stability of vanadate trapped [^3H]NANP on gizzard myosin was of considerable interest since specific photoincorporation of NANP into skeletal myosin has been shown to require previous trapping of some type^{1,2}. As Fig. 3 illustrates, both the gizzard myosin and the skeletal myosin-[^3H]NANP-vanadate complexes show essentially identical long-term stabilities. In both cases the half-life at 0°C was greater than 5 days indicating that leakage of [^3H]NANP from the active site would be negligible during the short time required for purification and photolysis of the trapped complex. It should be noted that care was taken to remove traces of actin from gizzard myosin and to keep the temperature near 0°C . Both factors, F-actin and elevated temperatures, are known to destabilize nucleotides trapped at the active site of skeletal myosin^{7,19}.

The gizzard myosin-[^3H]NANP-vanadate complex purified by gel filtration was photolysed as described in Fig. 4. Approxi-

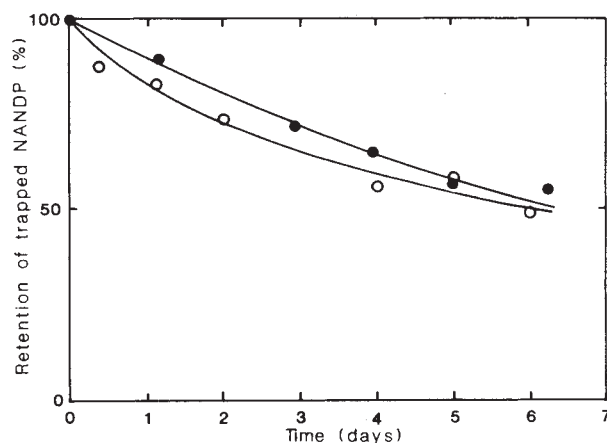


Fig. 3 Relative stability of [^3H]NANP-vanadate complexes of gizzard (○) and skeletal (●) myosins. The vanadate-[^3H]NANP complexes of gizzard and skeletal myosin were prepared and purified as described in Fig. 1 and stored at 0°C . At appropriate times, 0.1-ml aliquots were passed over PD-10 columns (Pharmacia) and analysed for radioactivity and protein concentration^{1,18}.

mately 0.20 mol per mol active site or 30% of the trapped [^3H]NANP (0.68 mol per mol active site) was found to be attached covalently after 40 min of photolysis. It was found to be important in the photolabelling to use irradiation above 400 nm as ultraviolet light caused a photochemical reaction of the trapped vanadate with myosin which in turn allowed [^3H]NANP (or in other experiments [^3H]ADP) to leak out of the active site (J.G., R.G.Y. and Y.O., unpublished results). In control experiments using the conditions in Fig. 4, less than 2% of either [^3H]NANP or [^3H]ADP trapped with vanadate on gizzard myosin was lost in 40 min of irradiation. Complete resolution of the heavy chains from the two light chains, LC_{20} (regulatory) and LC_{17} (essential), was accomplished using high-speed gel filtration on TSK-SW columns in the presence of SDS²⁰. Surprisingly, as Fig. 4 illustrates, some 20% of the photo-incorporated [^3H]NANP was associated with LC_{17} , with the remainder of the labelling occurring with the heavy chains. The LC_{20} chains were not labelled. These results suggest that parts of both LC_{17} and the heavy chain combine to form the purine-binding site for ATP. Such a finding is reminiscent of the structure of the immunoglobulin G molecule, in which parts of both the heavy chains and the light chains function to form the antigen binding site²¹. For IgG, as appears to be the case here, affinity labels gave the first indication that both chains were necessary to form the complete binding site^{22,23}.

The validity of the interpretation of our results is supported further by the remarkable similarity between NANP and ATP in their ability to support tension development and relaxation of chemically skinned muscle fibres from all three major types of muscle (R. Bridenbaugh, K. Nakamaye, G. Kerrick and R.Y., in preparation). In addition, support of effective tension development in smooth muscle by NANP required previous thiophosphorylation²⁴ of LC_{20} because NANP was not an effective substrate for myosin light-chain kinase. These results are clear evidence that NANP and, by analogy, NANP, binds to the active site of gizzard myosin in a manner similar if not identical to ATP.

The marked difference in the labelling pattern observed for smooth muscle myosin compared with that seen with striated muscle myosins is surprising. It is most likely that this difference reflects variations in the heavy-chain sequences because the sequences of the essential light chains from chicken gizzard, heart and skeletal muscle are all highly homologous²⁵. Currently, little is known about the sequence of smooth muscle myosin

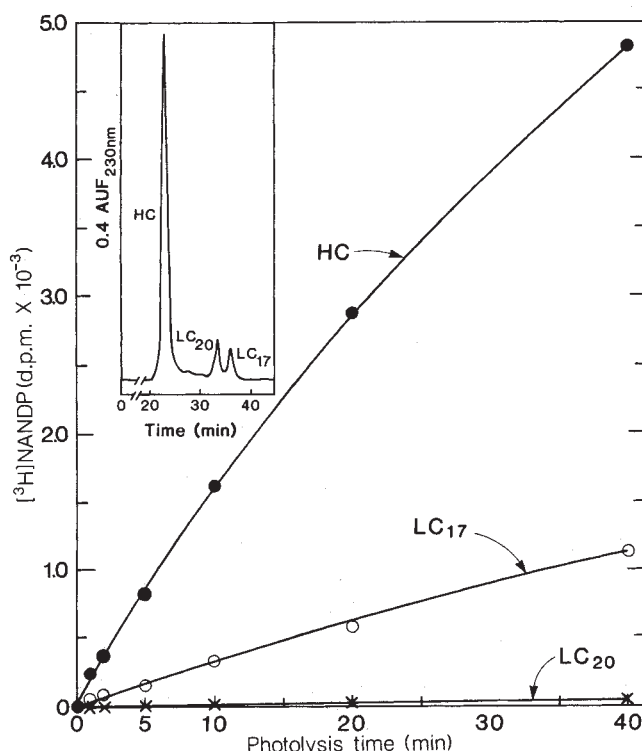


Fig. 4 The subunit location of $[^3\text{H}]$ NANDP photoincorporated into gizzard myosin. The gizzard myosin- $[^3\text{H}]$ NANDP-vanadate complex was prepared essentially as described in Fig. 1 except trapping was done overnight using 1 mM Na vanadate, 25 μM $[^3\text{H}]$ NANDP and 3.7 mg ml $^{-1}$ gizzard myosin. The Sephadex G-25 purified complex (10 ml) was irradiated in a 6 cm i.d. Petri dish at 0°C with a Ushio (100 W) high-pressure mercury lamp. The Petri dish was encased in aluminium foil and the top covered with a glass filter (L-42, Toshiba) to remove irradiation below 400 nm. At appropriate irradiation times, samples were taken from the Petri dish, cold trichloroacetic acid added to 5% and the protein precipitate collected by centrifugation. The precipitate was dissolved in Tris buffer and SDS (20% solution) added to 0.2%. The solution was heated at 100°C for 3 min, classified by centrifugation and 100 μl of the supernatant passed over serially linked TSK G4000SW and TSK G3000SW columns (7.5 mm $\times$ 60 cm each, Toyo Soda) using Beckman-Altex pumps at 100 p.s.i., flow rate = 1 ml per min. All three myosin polypeptides were well resolved (see inset) with the peak of the heavy chain (HC) eluting at 23 min, LC₂₀ at 33 min and LC₁₇ at 36 min. The counts were normalized at each time point to 1.4 mg of myosin (~ 6.0 nmol of active sites). The distribution of radioactivity was $78 \pm 4\%$ into HC and $17 \pm 2\%$ into LC₁₇.

heavy chains. It may be only in the case of myosin-linked regulated systems²⁶⁻²⁹ that the essential light chains function to provide part of the ATP binding site. Experiments to test this suggestion and to test the possible effect of phosphorylation of LC₂₀ and the folding of the heavy chain on the subunit labelling pattern of gizzard myosin are currently under way.

This research was supported in part by a research grant from the Ministry of Education, Science and Culture of Japan to Y.O. and T.S. and by NIH grant AM-05195 and by the Muscular Dystrophy Association to R.G.Y.

Received 11 June; accepted 5 August 1986.

- Okamoto, Y. & Yount, R. G. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1575-1579 (1985).
- Nakamaye, K., Wells, J. A., Bridenbaugh, R., Okamoto, Y. & Yount, R. G. *Biochemistry* **24**, 5226-5235 (1985).
- Wells, J. A. & Yount, R. G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4966-4970 (1979).
- Wagner, P. D. & Giniger, E. *Nature* **292**, 560-562 (1981).
- Sivaramakrishnan, M. & Burke, M. *J. biol. Chem.* **257**, 1102-1105 (1982).
- Goodno, C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2620-2624 (1979).
- Wells, J. A. & Yount, R. G. *Meth. Enzym.* **85**, 95-116 (1982).

- Mahmood, R. & Yount, R. G. *J. biol. Chem.* **259**, 12956-12959 (1984).
- Szilagyi, K., Balint, M., Sreter, F. A. & Gergely, J. *Biochem. biophys. Res. Commun.* **87**, 936-945 (1979).
- Adelstein, R. S. & Eisenberg, E. *Rev. Biochem.* **49**, 921-956 (1980).
- Marston, S. *Prog. Biophys. molec. Biol.* **41**, 1-41 (1982).
- Onishi, H. & Wakabayashi, T. *J. Biochem. (Tokyo)* **92**, 871-879 (1982).
- Trybus, K. M., Huiatt, T. W. & Lowey, S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6151-6155 (1982).
- Craig, R., Smith, R. & Kendrick-Jones, J. *Nature* **302**, 436-439 (1983).
- Suzuki, H., Onishi, H., Takahashi, K. & Watanabe, S. *J. Biochem. (Tokyo)* **84**, 1529-1542 (1978).
- Ikebe, M., Hinkins, S. & Hartshorne, D. J. *Biochemistry* **22**, 4580-4587 (1983).
- Trybus, K. M. & Lowey, S. *J. biol. Chem.* **259**, 8564-8571 (1984).
- Okamoto, Y. & Sekine, T. *J. Biochem. (Tokyo)* **87**, 167-178 (1980).
- Goodno, C. C. & Taylor, E. W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 21-25 (1982).
- Kato, Y., Komiya, K., Sasaki, H. & Hashimoto, T. *J. Chromatogr.* **193**, 29-36 (1980).
- Poljak, R. J. *Nature* **256**, 373-376 (1975).
- Singer, S. J. & Doolittle, R. F. *Science* **153**, 13-25 (1968).
- Good, A. H., Ovary, E. & Singer, S. J. *Biochemistry* **7**, 1304-1310 (1968).
- Cassidy, P., Hoar, P. E. & Kerrick, W. G. L. *J. biol. Chem.* **254**, 11148-11153 (1979).
- Matsuda, G. *Adv. Biophys.* **16**, 185-218 (1983).
- Adelstein, R. S. & Conti, M. A. *Nature* **256**, 597-598 (1975).
- Scordilis, S. P. & Adelstein, R. S. *Nature* **268**, 558-560 (1977).
- Trotter, J. A. & Adelstein, R. S. *J. biol. Chem.* **254**, 8781-8785 (1979).
- Szent-Gyorgyi, A. G., Szentkiralyi, E. M. & Kendrick-Jones, J. *J. molec. Biol.* **74**, 179-203 (1973).
- Ebashi, S. *J. Biochem.* **79**, 229-231 (1976).
- Goodno, C. *Meth. Enzym.* **85**, 116-123 (1982).

Convergent and divergent evolution of regulatory sites in eukaryotic phosphorylases

Peter K. Hwang & Robert J. Fletterick

Department of Biochemistry and Biophysics,
University of California, San Francisco, California 94143, USA

The activity of many proteins in eukaryotic cells is regulated by reversible covalent phosphorylation¹. This regulatory modification is often linked to other allosteric controls within the same protein², and such overlapping regulatory mechanisms are best characterized for glycogen phosphorylase (EC 2.4.1.1). Phosphorylases from different organisms or cell types exhibit markedly contrasting regulatory features³; this makes the enzyme attractive for studying the evolution of interacting molecular regulatory mechanisms^{4,5}. Extensive biochemical and crystallographic studies of rabbit muscle phosphorylase have led to a characterization of five regulatory regions (phosphorylation, glycogen storage, AMP, glucose and purine sites)⁶⁻⁸. Here we report the complete primary structure of the yeast *Saccharomyces cerevisiae* glycogen phosphorylase, deduced from the sequence of the cloned gene. Regions that are highly conserved between muscle and yeast enzymes include the active site, the glycogen storage site and possibly the glucose and purine inhibition sites. Partial conservation of the residues involved in AMP-binding suggests a binding site for the yeast enzyme inhibitor, glucose 6-phosphate^{9,10}. Other parts of the AMP site and the intersubunit contacts involved in AMP allostery are disrupted in the yeast enzyme by extreme sequence divergence. The poor alignment of amino termini and lack of homology at phosphorylation sites indicate that regulation by reversible phosphorylation evolved independently in yeast and vertebrate phosphorylases.

The yeast phosphorylase gene was isolated from a collection of yeast genomic DNA fragments cloned into plasmid vectors (see Fig. 1 legend). An 890-residue amino-acid sequence was deduced from the nucleotide sequence of a 2.7-kilobase (kb) uninterrupted open reading frame. We have compared the yeast amino-acid sequence to that of rabbit muscle phosphorylase (Fig. 2). The best alignment requires the introduction of 14 insertions or deletions in one sequence relative to the other (Fig. 2 legend). Although the alignment of N termini is somewhat arbitrary, as the two sequences differ considerably, the overall similarity between the yeast and muscle protein sequences is quite high. The yeast enzyme exhibits 49% amino-acid identity

Two of the five regulatory sites in rabbit muscle phosphorylase bind inhibitors. The enzyme is synergistically inhibited by glucose and purines¹¹, both of which bind in the active-site cleft formed between the N- and C-terminal domains¹²⁻¹⁴. Glucose binds directly at the active site loop (to Asn 284), which is conserved in the yeast enzyme (Fig. 3). Asp 283, which forms an important part of the loop interaction with Arg 569, is also conserved. Purines bind in a hydrophobic pocket formed by Phe 285 and Tyr 613; these residues are also conserved in the yeast sequence. Conservation of these glucose-binding and purine-binding residues suggests that glucose and purines are involved in the regulation of yeast phosphorylase although this has not previously been investigated.

Fig. 2 Comparison of the primary structures of yeast and rabbit muscle phosphorylases. The rabbit muscle phosphorylase sequence includes the corrections of Nakano *et al.*³³ and has been renumbered accordingly. The Dayhoff alignment program³⁴ was used to obtain the best homology. Further adjustment was by computer graphic modelling of the altered yeast side chains on the tertiary structure of the rabbit muscle enzyme. (Details of this procedure will be published elsewhere.) Gaps (-) have been inserted to maximize similarity. Only non-identical residues of rabbit muscle phosphorylase are indicated. At the N terminus, yeast phosphorylase is 28 residues longer than the muscle enzyme; at the C terminus, it is shorter by 11 residues. There are 12 other sites of insertions/deletions in the yeast sequence relative to muscle. Two deletions, of 8 and of 3 residues, are found at positions corresponding to loops in the rabbit muscle enzyme (316-323, 723-725). Of the 10 insertions, 5 involve additions of a single amino acid, preceding positions 438, 474, 735, 767 and 794 of the rabbit enzyme sequence. The rest are multiple-residue insertions: three residues at position 507, six at 213 and 557, nine at 591 and 13 at 114. Lower case letters, residues of rabbit muscle phosphorylase involved in catalytic or regulatory interactions: d, dimer or subunit interface; p, covalent phosphorylation site; a, AMP effector binding site; g, active site (and glucose inhibition); c, caffeine and purine inhibitor binding site; s, glycogen storage site; v, pyridoxal-phosphate cofactor binding site.

A more intricate control of glycogen metabolism is exerted by reversible covalent phosphorylation of phosphorylases. Although yeast and muscle phosphorylases interconvert between phosphorylated (active) and dephosphorylated (inactive) forms, many components of the mechanism are known to be different. For example, the phosphorylated peptide isolated from the N terminus of rabbit muscle phosphorylase is conserved in vertebrate phosphorylases¹⁷; the phosphorylated peptide in the yeast enzyme, however, does not match this sequence¹⁸. In addition, the specific modifying enzymes of muscle phosphorylase do not act on the yeast enzyme⁹. From the poor homology and different N terminus lengths (Fig. 2), it is obvious that the substrate sites for the interconverting enzymes are dissimilar. The target of phosphorylation in the yeast enzyme is a threonine (residue 19) that appears in the extended N terminus. The amino-acid sequence around this residue (Arg-Arg-Leu-Thr) matches the consensus sequence for phosphorylation sites in substrates of cyclic AMP-dependent protein kinase¹, in accordance with the finding that yeast phosphorylase can be phosphorylated by this enzyme^{19,20}. We cannot identify any homologous element in the



Fig. 3 Comparison of the three-dimensional structures of rabbit muscle and yeast phosphorylases at the active-site cleft. Darker segments of the α -carbon representation indicate identical or chemically similar residues conserved between the yeast and rabbit muscle enzymes; lighter segments indicate non-conserved residues. Predominantly in dark segments, this region appears structurally conserved in the yeast enzyme. Arrow, the active site, shown occupied by glucose and with conserved side chains Asp 283 and Asn 284. Side chains Phe 285 and Tyr 613 (thin lines) form the purine inhibition site (shown occupied by caffeine) and are also conserved in the yeast enzyme.

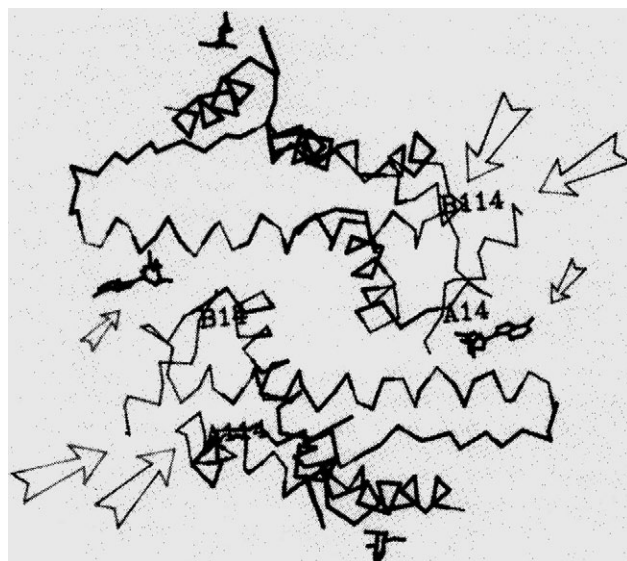


Fig. 4 Comparison of yeast phosphorylase to the three-dimensional structure of the rabbit muscle phosphorylase dimer (active conformation) at the AMP site, subunit interface and serine-14 phosphorylation site. Shown are projected α -carbon representations of a similar region in two contacting subunits (chains distinguished by labelling residues A and B). Identical or similar residues between rabbit muscle and yeast enzymes are indicated in darker segments. Lighter segments indicate non-conserved positions. Many differences occur in this region. The phosphorylation site, Ser 14 (positions A14 or B14), lies adjacent to the AMP site, shown occupied by AMP (small arrow). Larger arrows mark position 114, where yeast phosphorylase contains a 13-residue insertion, and the N terminus, where the yeast enzyme contains 28 extra residues.

yeast sequence with which the phosphorylation site of muscle phosphorylase (Ser 14) might align. In addition, there is no homology between the yeast N-terminal sequence and corresponding regions of other eukaryotic or prokaryotic phosphorylases^{5,21-25}.

The intersubunit contacts of the rabbit muscle phosphorylase dimer are also significant in the regulatory mechanisms of the enzyme^{6,7}. The structural segments that comprise the subunit interface couple the AMP and phosphorylation sites to the active-site region^{6,26}. Major differences between muscle and yeast enzymes occur in this region (Fig. 4). We have already noted specific alterations at the AMP binding site and the phosphorylation site, and the lack of homology at the N terminus. Other residues involved in the intersubunit contact of muscle phosphorylase are indicated in Fig. 2; most of these are not conserved in the yeast sequence. In addition, there are two major length variations at positions near the dimer interface; at the N terminus, yeast phosphorylase has an extra 28 residues compared with the muscle homologue, and at a position homologous to residue 113 of the muscle enzyme, the yeast enzyme contains an insertion of 13 residues. Thus, the subunit interface, as found in rabbit muscle phosphorylase, appears to be altered in the yeast enzyme, implying that intersubunit cooperativity in yeast phosphorylase involves structural components and mechanisms that are distinct from those in the muscle enzyme.

Yeast phosphorylase is strongly inhibited by glucose 6-phosphate^{9,10}. Although our analysis shows that a functional AMP activator site is disrupted in the yeast enzyme, conservation of the phosphoryl subunit and partial conservation of the sugar subunit suggest that this region serves as the binding site for glucose 6-phosphate. This hypothesis is supported by the following evidence. Inhibition of yeast phosphorylase by glucose 6-phosphate is non-competitive with respect to both glucose 1-phosphate and glycogen⁹, implying that the effector binds somewhere other than at the catalytic site. Glucose 6-phosphate does

bind at the AMP site of inactive (dephosphorylated) muscle phosphorylase and inhibits activation (phosphorylation)^{16,27}. Crystallographic studies¹⁶ have implicated all the conserved residues between yeast and muscle enzymes at the AMP site in the binding of glucose 6-phosphate to rabbit muscle phosphorylase. Trp 67, another residue in contact with glucose 6-phosphate in rabbit muscle phosphorylase¹⁶, is also conserved in the yeast enzyme.

The lack of sequence homology at the N terminus indicates that the phosphorylation sites of yeast and vertebrate phosphorylases did not evolve from a common ancestral regulatory region. Palm *et al.*⁵ have suggested that a preformed nucleotide sequence encoding a phosphorylatable peptide was fused to a primordial phosphorylase gene in the evolution of vertebrate phosphorylases. If true, it seems that a separate gene fusion event generated the phosphorylation site in yeast phosphorylase. Of course, the acquisition of a phosphorylation site alone would not have been sufficient for covalent regulation, because allosteric regulation by reversible phosphorylation requires intersubunit interactions that couple phosphorylation to active-site changes. The N-terminal sequence divergence that generated different phosphorylation sites in yeast and muscle enzymes also affected development of appropriate intersubunit contacts. Although its structure remains to be elucidated, the subunit interface in yeast phosphorylase must differ significantly from that of muscle phosphorylase but still functions analogously, transmitting regulatory responses to the catalytic site. Hence, the intersubunit interactions responsible for allosteric transmission in yeast and muscle phosphorylases seem to have arisen by convergent evolution. Presumably, the dissimilar N termini also influenced the completion of the coupled allosteric ligand-binding site, enabling development of contrasting regulatory functions, glucose 6-phosphate inhibition in yeast and AMP activation in muscle.

Finally, how has the regulatory phosphorylation cascade evolved with time? The phosphorylated and dephosphorylated forms

of muscle phosphorylase are interconverted by phosphorylase kinase and phosphorylase phosphatase, specific enzymes that, in turn, are regulated (phosphorylated) by cyclic AMP-dependent protein kinase. The phosphorylation site of yeast phosphorylase, however, is recognized by both the cyclic AMP-dependent protein kinase and a specific phosphorylase kinase^{19,20}, as well as by a general protein phosphatase^{28,29}. The following evolutionary model is consistent with these observations: a phosphorylation site recognized by a general protein kinase (and phosphatase) was initially spliced onto the N terminus of the ancestral enzyme; divergence of this site led to poorer recognition and eventually loss of recognition by the general protein kinase; coevolution of other kinases (and phosphatases) gradually restricted their range of substrates, resulting in specificity for phosphorylase. The difference in phosphorylation site specificity between yeast and muscle phosphorylases may reflect how far the regulatory cascades of yeast and vertebrates have evolved in this general path.

We thank Drs Michael Stern, Susan Michaelis, Christopher Newgard, Kenichi Nakano and John Burke for helpful advice, discussions and technical assistance. Supported by a grant from the NSF (PCM84-20489).

Received 19 May; accepted 7 August 1986.

1. Cohen, P. *Eur. J. Biochem.* **151**, 439-448 (1985).
2. Cohen, P. *Control of Enzyme Activity* 2nd edn, 1-96 (Chapman & Hall, London, 1983).

3. Graves, D. J. & Wang, T. H. in *The Enzymes* Vol. 7, 3rd edn (ed. Boyer, P.), 430-482 (Academic, London, 1982).
4. Fischer, E. H., Pocker, A. & Saari, J. C. *Essays Biochem.* **6**, 23-68 (1970).
5. Palm, D., Goerl, R. & Burger, K. *J. Nature* **313**, 500-502 (1985).
6. Fletterick, R. J. & Sprang, S. R. *Acc. chem. Res.* **15**, 361-369 (1983).
7. Fletterick, R. J. & Madsen, N. B. A. *Rev. Biochem.* **49**, 831-861 (1980).
8. Dombradi, V. *Int. J. Biochem.* **13**, 125-139 (1981).
9. Fosset, M., Muir, L. W., Nielsen, L. D. & Fischer, E. H. *Biochemistry* **10**, 4105-4113 (1971).
10. Sagardia, F., Gotay, I. & Rodriguez, M. *Biochem. biophys. Res. Comm.* **42**, 829-935 (1971).
11. Kasvinsky, P. J., Schechosky, S. & Fletterick, R. J. *J. Biol. Chem.* **253**, 9102-9106 (1978).
12. Guenard, D., Morange, M. & Buc, H. *Eur. J. Biochem.* **76**, 447-452 (1977).
13. Stura, E. A. *et al. J. molec. Biol.* **170**, 529-565 (1983).
14. Sprang, S. *et al. Biochemistry* **21**, 2036-2048 (1982).
15. Sprang, S. R. & Fletterick, R. J. *J. molec. Biol.* **131**, 523-551 (1979).
16. Lorek, A. *et al. Biochem. J.* **218**, 45-60 (1984).
17. Titani, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 4762-4766 (1977).
18. Lerch, K. & Fischer, E. H. *Biochemistry* **14**, 2009-2014 (1975).
19. Becker, J.-U., Wingender-Drissen, R. & Schiltz, E. *Archs Biochem. Biophys.* **225**, 667-678 (1983).
20. Pohlig, G., Wingender-Drissen, R. & Becker, J.-U. *Biochem. biophys. Res. Commun.* **114**, 331-338 (1983).
21. Cohen, P., Saari, J. C. & Fischer, E. H. *Biochemistry* **12**, 5233-5241 (1973).
22. Schiltz, E., Palm, D. & Klein, H. W. *FEBS Lett.* **109**, 59-62 (1980).
23. Lacks, S. A., Dunn, J. J. & Greenberg, B. *Cell* **31**, 327-336 (1982).
24. Nakano, K., Fukui, T. & Matsubara, H. *J. Biol. Chem.* **255**, 9255-9261 (1980).
25. Newgard, C. B., Nakano, K., Hwang, P. K. & Fletterick, R. J. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
26. Sprang, S. & Fletterick, R. J. *Biophys. J.* **32**, 175-192 (1980).
27. Madsen, N. B. & Schechosky, S. *J. Biol. Chem.* **242**, 3301-3307 (1967).
28. Wingender-Drissen, R. & Becker, J.-U. *Biochim. biophys. Acta* **743**, 343-350 (1983).
29. Matsumoto, K., Uno, I., Kato, K. & Ishikawa, T. *Yeast* **1**, 25-38 (1985).
30. Carlson, M. & Botstein, D. *Cell* **28**, 145-154 (1982).
31. Southern, E. J. *J. molec. Biol.* **98**, 503-517 (1975).
32. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161-178 (1980).
33. Nakano, K., Hwang, P. K. & Fletterick, R. J. *FEBS Lett.* (in the press).
34. Dayhoff, M. O. (ed.) in *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 2, 3-8 (Natl. Biomed. Res. Fdn, Silver Spring, Maryland, 1976).

α -Lactalbumin possesses a novel calcium binding loop

D. I. Stuart*, K. R. Acharya*, N. P. C. Walker†‡, S. G. Smith*, M. Lewis*‡ & D. C. Phillips*

* Laboratory of Molecular Biophysics, The Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK

† Chemistry Department, Queen Mary College, Mile End Road, London E1 4NS, UK

Calcium performs a unique role in biology, achieving biological effects through highly specific interactions with and modulation of target proteins¹. It has been proposed that calcium-modulated proteins possess a characteristic, evolutionarily related, binding fold, known as the EF-hand². The high-resolution X-ray structure of α -lactalbumin reveals a Ca^{2+} binding fold that resembles an EF-hand only superficially and presumably has no evolutionary relationship with it. However, there is clear homology with the corresponding loop in c-type lysozyme (the 'parent' molecule of α -lactalbumin). This study, at 1.7 Å resolution, represents one of the most accurate analyses of a calcium binding protein yet reported.

α -Lactalbumin acts as a specifier protein³, it comprises some 15% of total protein in human milk⁴ and is essential for the production of lactose. The attachment of α -lactalbumin to β -galactosyl transferase on the luminal surface of the Golgi membrane results in a complex, lactose synthetase that catalyses the addition of galactose to glucose to produce lactose. α -Lactalbumin is a small protein that has about 40% sequence identity with c-type lysozyme, and also a similar intron-exon genetic organization⁵, strongly suggesting that the two have diverged from a common ancestor. α -Lactalbumin is a calcium metalloprotein⁶, the K_{app} in physiological conditions is likely to be of the order of 10^{-6} - 10^{-7} M (ref. 7), comparable to that observed for the calcium-modulated proteins¹. In addition, Zn^{2+} binding

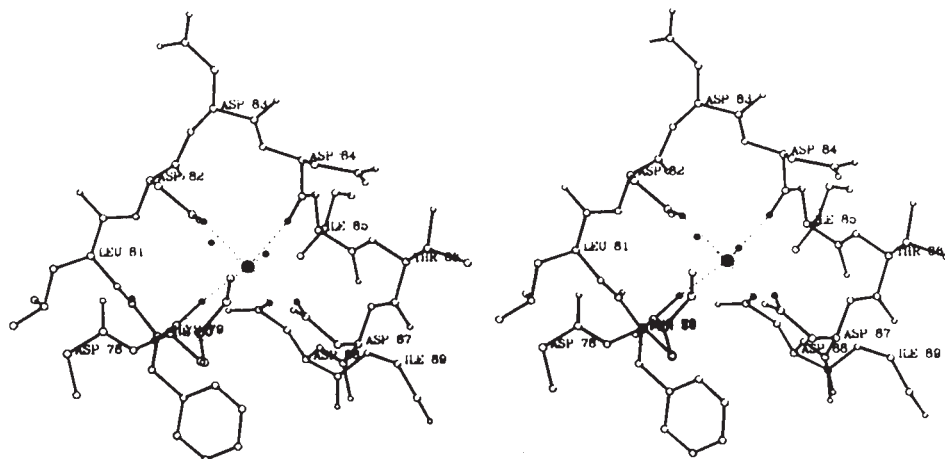
has been observed (K_d 5 μM)⁸. The biological role of metal ion binding is unclear. If the concentration of free calcium ions in the Golgi lumen is comparable to that in milk (about 3 mM), it would be expected that the Ca^{2+} site would be fully occupied, whereas the occupation of the Zn^{2+} site would be uncertain. The first role proposed for the Ca^{2+} binding was to stabilize the protein and protect it from thermal denaturation^{6,9}; more recently it has been claimed that the pattern of ion binding affects the catalytic properties (notably the V_{max}) of the lactose synthetase complex⁸. Calcium binding may also be important in controlling the release of α -lactalbumin from the Golgi membrane, an essential step in inducing lactation⁸. It is plausible that these mechanisms allow fine control of lactose synthesis.

The second messenger role of Ca^{2+} ions, binding to target proteins and hence inducing a conformational change, is well established. Unfortunately, in no case is the precise mechanism of triggering understood in terms of the concomitant conformational changes, although the crystal structure of troponin-C gives some hints¹⁰. In the case of α -lactalbumin, knowing the structure of lysozyme, a close homologue with no marked propensity for binding calcium, it is possible to infer something of the conformational changes induced by the calcium binding by comparison with the structure of the calcium-bound form of α -lactalbumin.

Structural studies on crystalline α -lactalbumins have been in progress for a number of years but have been frustrated by various problems^{11,12}. We have determined the structure of α -lactalbumin from baboon milk. Low-resolution isomorphous replacement phases allowed the approximate orientation of the molecule in the unit cell to be determined and confirmed that the structure is broadly similar to lysozyme^{13,14}. A process of careful refinement (initially of a lysozyme model) against the X-ray observations revealed the detailed structure. This work will be reported in detail elsewhere. The sequence of the enzyme from baboon milk has not been determined, but preliminary experiments suggested very few differences between the sequences of human and baboon α -lactalbumins (R. Greenberg, personal communication). Our present model therefore has few changes from the human α -lactalbumin sequence. The R -factor³² of 0.207 (for 10,627 data $>2\sigma(F)$) to a resolution of 1.7 Å together with the reasonable stereochemistry exhibited by the model (r.m.s. (root mean square) deviation from ideal

‡ Present addresses: BASF AG, Hauptlaboratorium A30, D-6700 Ludwigshafen, FRG (N.P.C.W.); Smith Kline and French Laboratories, 1500 Spring Garden Street, P.O. Box 7929, Philadelphia, Pennsylvania 19101, USA (M.L.).

Fig. 1 Stereoscopic views of the calcium binding loop in α -lactalbumin. Residues 78–89 plus the calcium ion (larger black dot) and its water ligands. Dashed lines join the calcium to its ligands.



covalent bond lengths = 0.03 Å) and the high quality of the electron density map lead us to be confident of our interpretation of the molecular structure. We have placed all the protein atoms and 132 water molecules. We estimate that the coordinates for the parts of the structure that we discuss below are accurate to within a few tenths of an ångström unit. The determination of the α -lactalbumin structure represents a notable application of the molecular replacement method as the starting model has only 37% sequence identity with α -lactalbumin and is a distinct enzyme.

When the enzyme is crystallized and no metals added to the crystallization medium, as expected from the binding affinity data, the calcium binding site is occupied. Nucleic acid sequences¹⁵ and revised amino-acid sequences¹⁶ reveal two pairs of strongly conserved aspartic acid residues (82, 83 and 87, 88). This led to the proposal¹⁶, subsequently supported by NMR data¹⁷, that these four residues are concerned in calcium binding. We find that three of these residues are calcium ligands. The Ca^{2+} site is easily identifiable as a persistent feature of our difference Fourier maps which cannot be adequately accounted for by a water molecule. Refinement fully confirms this assignment, the current B -values for the ion and its ligands being about 15 Å² and indeed the B -factors for all this region are among the lowest in the molecule. Figure 1 shows the detailed atomic structure around the calcium, which is surrounded by seven ligands, all oxygens, three from the carboxylate groups of aspartyl residues (82, 87 and 88), two carbonyl oxygen atoms (from residues 79 and 84) and two water molecules. The five ligands contributed by the protein come from a rather tight turn akin to a β -turn with an extra residue, Asp 82, acting as a 'spacer' between two helices. We shall subsequently refer to this structure as the 'elbow'. The geometry in small-molecule- Ca^{2+} complexes is variable but generally involves six, seven or eight ligands with a distorted pentagonal bipyramid being the most common configuration in the seven coordinate case (G. Klebe, personal communication). When calcium binds to proteins, the solvent often contributes one or more ligand¹⁸. In α -lactalbumin the coordination is ideal, forming a slightly distorted pentagonal bipyramid with carbonyl oxygen atoms at the apices while the carboxylate groups and water are arranged with the waters approximately opposed (Fig. 1). The Ca–O distances, ranging from 2.3 Å to 2.5 Å (Table 1), are close to the expected value of 2.4 Å.

The only comparable calcium binding sites described for other proteins (in terms of affinity and topology) are those forming the class of folds known as the EF-hand². Comparison of the α -lactalbumin loop with an EF-hand (see Fig. 2) reveals that although the elbow of α -lactalbumin is a more compact structure (consisting of two fewer residues) there are certain marked similarities; both structures are formed from a loop between almost orthogonal helices and deploy similar types of ligands around the Ca^{2+} . In α -lactalbumin a single Asp turn¹⁹ stabilizes

the structure whilst the longer EF-hand uses several of these. Furthermore, in both cases there is a conserved central hydrophobic residue (Ile 85 in α -lactalbumin) which is important in stabilizing the structure as a whole. We can discount the naive hypothesis of an homologous relationship for several reasons. Firstly because the evolutionary history of α -lactalbumin rules out such a scheme. Secondly, the Ca^{2+} coordination observed for the typical EF-hand is considered a distorted octahedron compared with the pentagonal bipyramid of α -lactalbumin. Finally, on searching a database of well-refined structures for related conformations²⁰, the only structures found that were closely similar to a given EF-hand were other EF-hands. These structures superimpose to better than 1 Å r.m.s. deviation. As there are a large number of unrelated structures that superimpose to within about 2 Å, there is no evidence for significance in the structural similarity with the elbow where the r.m.s. deviation exceeds 2 Å.

Comparing the calcium binding loop in α -lactalbumin with the corresponding structure in hen egg white lysozyme we find, surprisingly, that the conformation of the backbone is essentially unchanged whereas the side chains are radically altered (Fig. 2). The r.m.s. deviation in C_α positions for this loop is 1.2 Å. The differences are largely attributable to movement of the loop away from the body of the molecule in α -lactalbumin compared with lysozyme. If we superimpose the coordinates of the loop alone, we find that they agree with an r.m.s. deviation of 0.6 Å, as close a similarity as that found between the calcium binding loops of the EF-hand proteins¹⁸ and indeed as that between human and hen lysozyme. Further evidence of the high degree of similarity comes from an analysis of the water structure around this loop in α -lactalbumin and lysozyme. There is a network of three water molecules in hen egg white lysozyme that corresponds to two water molecules and the calcium ion in α -lactalbumin²¹. The fold at the elbow seems to be unusual; no other loops superimpose well. The best fit (with rhodanese, residue 278 onwards) has an r.m.s. deviation of 1.8 Å and many others are about as good. Hence, although this loop is a region of considerable similarity in structure, there are important differences in sequence (Table 2).

Table 1 The calcium ligands

Residue	Group		Distance (Å)
79	Lys	Carbonyl O	2.3
82	Asp	Carboxylate OD1	2.4
84	Asp	Carbonyl O	2.3
87	Asp	Carboxylate OD1	2.3
88	Asp	Carboxylate OD1	2.4
		Water O	2.4
		Water O	2.5

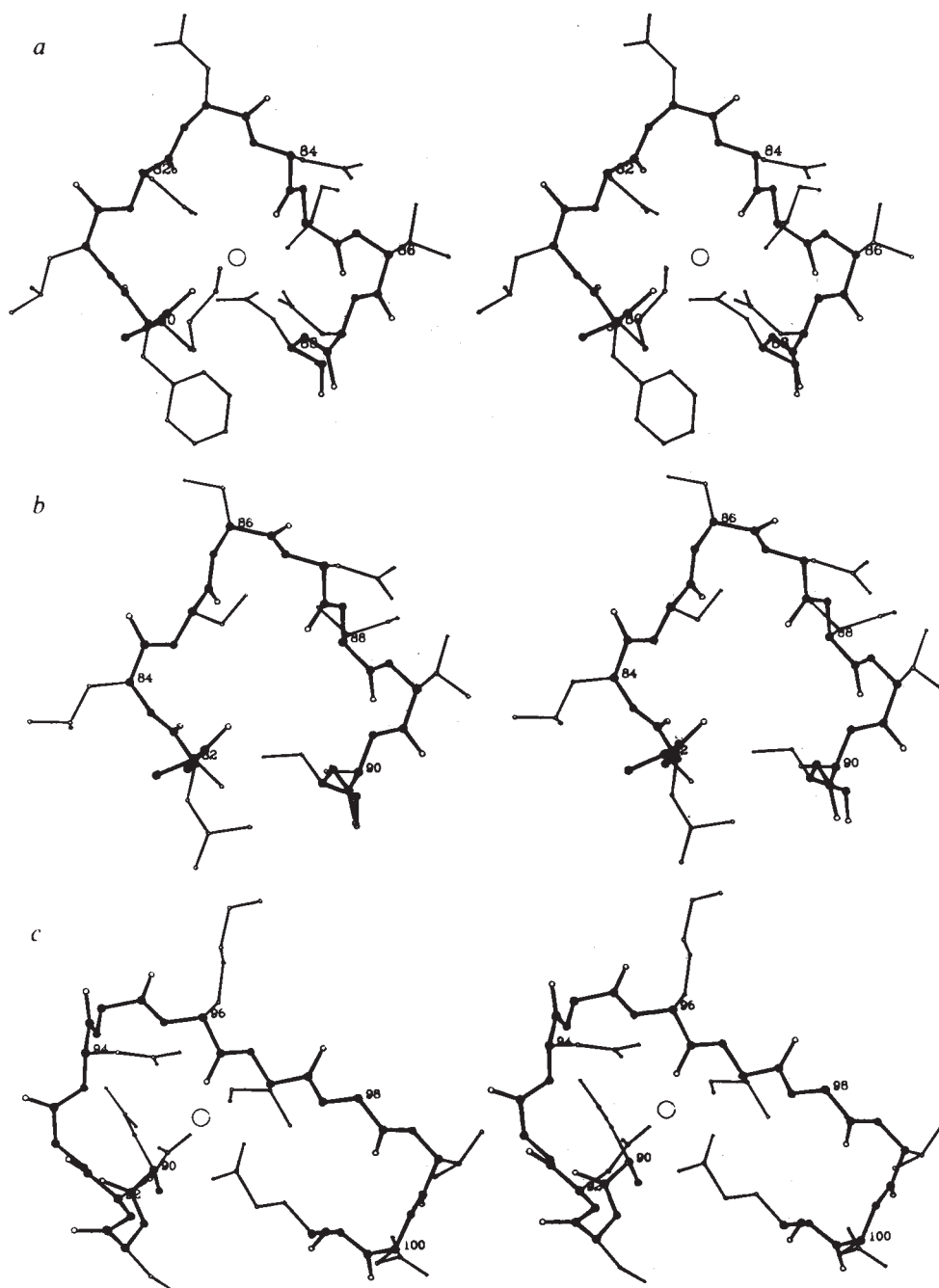


Fig. 2 Rotationally aligned structures of *a*, residues 79-88 in α -lactalbumin, *b*, residues 82-91 in hen lysozyme, *c*, residues 90-101 in carp parvalbumin, the EF-hand, data set F6A in the protein data bank²⁹. The central, enveloped atom in *a* and *c* is calcium. The superimpositions were performed using a computer program³⁰ drawing on the ideas of Rossmann and Argos³¹. The optimal rotation matrices were determined on the basis of the C_{α} coordinates of the residues in the loops (all 10 residues of the α -lactalbumin loop were matched with parvalbumin residues to give an r.m.s. deviation of 2.5 Å).

The above analysis refers to the hen lysozyme structure. We have also looked at human lysozyme²² where the three aspartates of α -lactalbumin that contribute carboxylate ligands correspond to Gln 86, Asp 91 and Ala 92. Asp 91 is rotated by $\sim 90^\circ$ about χ_1 (compared to α -lactalbumin), to bring the charged atoms into solution, whereas the ϵ -nitrogen of Gln 86 is at the position of a water ligand of Ca^{2+} in α -lactalbumin. It seems therefore that human lysozyme contains a partially formed calcium binding site without a destabilization of protein structure. Ala 92 is rather internal and its conversion to an Asp in the absence of some compensating charge or change in conformation would be expected to destabilize the protein. A survey of lysozyme sequences reveals that the sequence of horse milk lysozyme²³ shows the full complement of the three aspartate residues that form Ca^{2+} ligands in α -lactalbumin. If this sequence is correct, it raises the intriguing possibility that horse milk lysozyme might bind calcium.

It has been suggested that the entropic effects of calcium binding are more unfavourable in α -lactalbumin than the EF-

hand proteins²⁴. At first sight this is puzzling; the B -values for this loop in lysozyme are low²¹, and therefore entropic effects due to ordering of the loop on calcium binding should be small. Furthermore the rigidity of the calcium binding loop in α -lactalbumin is responsible for the discrimination the site shows between Ca^{2+} and Mg^{2+} (the K_d s differ by approximately 10^5), the loop being unable to collapse around the smaller Mg^{2+} ion. Nevertheless, uncompensated negative charges should destabilize the observed lysozyme conformation, so what conformational changes are likely? The results of the database search do not indicate any closely related folds that could be formed by rearrangement of the loop but a variety of structures are observed that differ by about 2 Å. No quantitative evidence is available on the Ca^{2+} -induced conformational changes²⁵ but a possible mode of action could involve a relatively small conformational change at the Ca^{2+} binding site, perhaps triggering rigid body movements of secondary structural units as has been described for other structural transitions²⁶.

The observations on the structure of the calcium binding

Table 2 Amino-acid sequences at the α -lactalbumin elbow

Protein	79	80	81	82	83	84	85	86	87	88
α -lac, human	Lys	Phe	Leu	Asp	Asp	Asp	Ile	Thr	Asp	Asp
α -lac, bovine	Lys	Phe	Leu	Asp	Asp	Asp	Leu	Thr	Asp	Asp
Lys, hen	Ala	Leu	Leu	Ser	Ser	Asp	Ile	Thr	Ala	Ser
Lys, human	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Ala
Lys, horse	Lys	Leu	Leu	Asp	Glu	Asn	Ile	Asp	Asp	Asp
Calcium ligands in α -lac	+			*		+			*	*

The residue numbering is based on human α -lactalbumin. Add three to these numbers to get the classical hen lysozyme numbering.

* Residues that contribute a side-chain calcium ligand; + residues that contribute a carbonyl oxygen ligand.

elbow are relevant to protein evolution. The divergence of α -lactalbumin from lysozyme involves the formation of a novel enzyme system, and seems to be related to the formation of an extremely well developed Ca^{2+} binding ability, a surprising innovation. We would argue against an alternative evolutionary interpretation, that an ancestral lysozyme possessed Ca^{2+} binding activity and the current lysozyme structure retains a 'memory' of this, because the proposed serial homologues (phage, goose and hen type lysozymes²⁷) do not possess this loop as a common feature. It is an interesting coincidence, however, that the loop of phage lysozyme once proposed as a calcium binding loop is probably homologous to a region fairly close to the elbow in α -lactalbumin^{27,28}. Clearly the evolution of a novel functional property of the α -lactalbumin molecule has not come about through exon shuffling or any sophisticated structural modifications. It seems instead to be a structurally localized property that arose principally through simple amino-acid substitutions at three positions to produce a calcium binding site of exquisite specificity. While other cases of divergence of structures to fulfil radically different functions are known (for example, the various functionally diverse immunoglobulin domains), here the process is so recent that we feel that we have almost caught evolution *flagrante delicto*. Naturally α -lactalbumin may lend support to the rather simplistic approach to protein engineering that considers function to be a local property. Such a hypothesis can be tested easily in this case by investigating the effects of mutagenic modification of the lysozyme loop in an attempt to produce a calcium binding site (as may have already been accomplished naturally in the horse milk lysozyme). We would be pleased if our observations stimulate such studies and studies on the biological significance of calcium binding to the lactose synthetase. Although the lactose synthetase complex is not cytosolic, it is interesting to see essentially intracellular calcium binding at a site with no homology to the EF-hand. At any event it is encouraging to see nature innovating rather than restricting itself to spare-functional-part protein design.

This work was supported by the MRC and SERC.

Received 14 April; accepted 18 August 1986.

- Levine, B. A. & Williams, R. J. P. in *Calcium and Cell Function*, Vol. II (ed. Cheung, W. Y.) 1-38 (Academic, New York, 1982).
- Kretsinger, R. H. & Nockolds, C. E. *J. biol. Chem.* **248**, 3313-3326 (1973).
- Brew, K. & Campbell, P. N. *Biochem. J.* **102**, 258-269 (1967).
- Aschaffenburg, R. & Drewry, J. *Biochem. J.* **65**, 273-277 (1957).
- Qasba, P. K. & Safaya, S. K. *Nature* **308**, 377-380 (1984).
- Hiraoka, Y. *et al. Biochem. biophys. Res. Commun.* **95**, 1098-1104 (1980).
- Mitani, M. *et al. J. biol. Chem.* **261**, 8824-8829 (1986).
- Musci, G. & Berliner, L. J. *Biochemistry* **24**, 6945-6948 (1985).
- Hiraoka, Y. & Sugai, S. *Int. J. Pept. Prot. Res.* **23**, 535-542 (1984).
- Herzberg, O., Moulty, J. & James, M. N. G. *J. biol. Chem.* **261**, 2638-2644 (1986).
- Aschaffenburg, R. *et al. J. molec. Biol.* **67**, 525-528 (1972).
- Fenna, R. E. *J. molec. Biol.* **161**, 203-210 (1982).
- Aschaffenburg, R. *et al. J. molec. Biol.* **127**, 135-137 (1979).
- Smith, S. G. *et al. Biochem. J.* (submitted).
- Hall, L. *et al. Nucleic Acid Res.* **10**, 3503-3515 (1982).
- Shewale, J. G., Sinha, S. K. & Bew, K. J. *biol. Chem.* **259**, 4947-4956 (1984).
- Kuwajima, K., Hiroshima, Y. & Sugai, S. *Int. J. Pept. Prot. Res.* **27**, 18-27 (1986).
- Herzberg, O. & James, M. N. G. *Biochemistry* **24**, 5298-5302 (1985).
- Rees, D. C., Lewis, M. & Lipscomb, W. N. *J. molec. Biol.* **168**, 367-387 (1983).
- Jones, T. A. & Thirup, S. *EMBO J.* **5**, 819-822 (1986).
- Handoll, H. D. *thesis*, Univ. Oxford (1985).
- Artymiuk, P. J. & Blake, C. C. F. *J. molec. Biol.* **152**, 737-762 (1981).

- McKenzie, H. A. & Shaw, D. C. *Biochem. Int.* **10**, 23-31 (1985).
- Schaer, J.-J., Milos, M. & Cox, J. A. *FEBS Lett.* **190**, 77-80 (1985).
- Koga, K. & Berliner, L. J. *Biochemistry* **24**, 7257-7262 (1985).
- Chothia, C. & Lesk, A. M. *Trends biochem. Sci.* **11**, 116-118 (1985).
- Weaver, L. H. *et al. J. molec. Evol.* **21**, 97-111 (1985).
- Tufty, R. M. & Kretsinger, R. H. *Science* **187**, 167-169 (1975).
- Bernstein, F. C. *et al. J. molec. Biol.* **112**, 535-542 (1977).
- Stuart, D. I. *thesis*, Bristol Univ. (1979).
- Rossmann, M. G. & Argos, P. *J. molec. Biol.* **109**, 99-129 (1977).
- Blundell, T. L. & Johnson, L. N. in *Protein Crystallography*, 147 (Academic, New York, 1976).

Bent DNA at a yeast autonomously replicating sequence

Michael Snyder*†, Andrew R. Buchman‡
& Ronald W. Davis†

* Department of Biology, Yale University, New Haven, Connecticut 06511, USA

† Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA

‡ Department of Cell Biology, Stanford University, Stanford, California 94305, USA

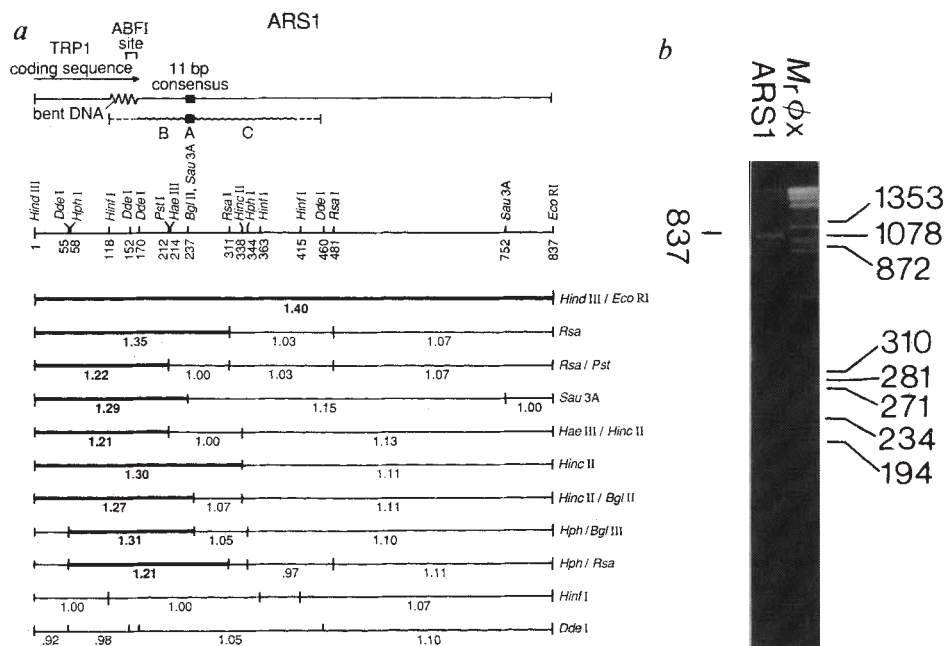
DNA fragments that show retarded electrophoretic mobility through polyacrylamide gels have been found in both prokaryotes and eukaryotes¹⁻⁷. In the case of kinetoplast DNA, evidence has been presented that the DNA is curved or 'bent'^{8,9}. Bent DNA has previously been found at the λ and simian virus 40 (SV40) DNA replication origins^{6,7}. Here we show the existence of bent DNA at a yeast autonomously replicating sequence (ARS1), a putative replication origin. The bent DNA has been localized to a 40-55 base pair (bp) segment and contains six (A)₃₋₅ stretches (that is, six poly(A) stretches, three to five nucleotides in length) phased approximately every 10.5 bp. This region contains a DNA binding site for a yeast protein factor. This site lies at the 3' end of the *TRP1* gene, in a region devoid of nucleosomes, and is positioned 80 bp away from the ARS consensus sequence; removal of this region impairs ARS function *in vivo*. The bent DNA may be involved in transcription termination or the prevention of nucleosome assembly in this region.

Yeast ARS sequences have properties expected of origins of replication: they allow the autonomous replication of sequences to which they are attached, occur at a frequency expected of chromosomal origins, replicate once per cell cycle using genes responsible for chromosomal replication and act as origins of replication *in vitro* (see ref. 10 for review). Electron microscopic data for ribosomal DNA and 2 μ DNA has mapped nascent replication bubbles near ARS elements suggesting that ARSs act as replication origins *in vivo*^{11,12}.

The structure of ARS1, which is near the 3' end of the yeast *TRP1* gene, is shown in Fig. 1. There are three elements that comprise ARS1, domains A, B and C. Domain A is an 11-bp consensus sequence found at most other ARS sequences.^{13,14} Mutations in this sequence at ARS1 or the ARS at the HO locus obliterate ARS function *in vivo*^{13,15,16}. Domains B and C flank ARS1 and are important for its function¹⁶⁻¹⁸; mutations in either

Fig. 1 *a*, Features of yeast ARS1; *b*, mobility of ARS1 restriction fragments on a 10% polyacrylamide gel.

A (box), B and C (— lines) are indicated. The broken lines indicate the degree of uncertainty of the extent of domains B and C. The *Hind*III/*Eco*RI ARS1 fragment was digested with the various restriction enzymes indicated, ethanol-precipitated and the electrophoretic mobilities of the fragments measured in a 10% polyacrylamide gel using 50 mM Tris base, 50 mM boric acid, 1 mM EDTA buffer²⁰. Electrophoresis was at 2.9 V cm⁻¹ for 16 h at 24.5 °C. Correct assignment of the DNA bands was determined by digestion with a second enzyme. Beneath the fragments is indicated the *k* value = expected distance migrated/observed migration distance. Fragments where *k* > 1.20 are indicated in bold type (not determined for fragments less than 60 bp). *b*, Migration of the 837 bp (determined from DNA sequence²³) *Hind*III/*Eco*RI ARS1 restriction fragment in a 10% polyacrylamide gel. 837 indicates the position of the 837-bp ARS1 fragment. Sizes of Φ X174 *Hae*III marker fragments are given. The 281-bp fragment of Φ X174 migrates more slowly than expected. Other standards used are λ *Bst*EII markers and pBR322 *Hin*FI markers (not shown).



region reduce, but do not abolish, ARS function *in vivo*. Domain B begins very close to the 3' end of the *TRP1* gene and may overlap it. A protein factor called ABFI (ARS binding factor I) that has been partially purified from yeast extracts,²⁰ binds in this region. The roles of the different ARS domains A, B and C, and of ABFI are unknown.

Because bent DNA has been found at bacteriophage and viral DNA replication origins^{6,7}, we tested for unusual DNA conformational features in the ARS1 region by subjecting restriction fragments of ARS1 DNA to electrophoresis in 10% polyacrylamide gels and identifying DNA fragments that exhibited unusually retarded mobility^{19,21,22}. The mobilities were compared to 'normal' DNA fragments derived from pBR322, Φ X174 and λ DNAs (see Fig. 1 legend). Several DNA fragments were found that migrate 20–40% more slowly than predicted from the known DNA sequence²³ at 24.5 °C (Fig. 1). For example, an 837-bp fragment migrates to a position expected of a 1,170-bp fragment, 40% more slowly than anticipated. Using the method of Diekmann and Wang¹⁹, several restriction fragments spanning the ARS1 region were tested for abnormal electrophoretic mobility and their *k* values determined (*k* = expected distance migrated/observed distance migrated for the DNA fragments). All DNA fragments that contain the 150-bp region between the *Hph*I and *Pst*I sites are found to have *k* values > 1.20. DNA fragments generated by cleaving the DNA with *Hin*FI and *Dde*I that cut within this 150-bp segment do not show abnormal mobility. Data presented below demonstrate that the *Hin*FI (118) and *Dde*I (152, 170) sites bracket a region which has properties characteristic of bent DNA. It is clear that there is a strong correlation between the position of the 118–170 *Hin*FI/*Dde*I region within a fragment and the observed *k* value. It has been noted that when the bent sequence is near the end of a fragment, the effect on retarded mobility is diminished²¹. Similar results were obtained after the digested ARS1 DNA was extracted with phenol or treated with proteinase K (data not shown). Thus reduced mobility does not appear to be due to association with proteins.

The abnormal DNA region was further localized by deletion analysis. A series of five deletions beginning at the *Hind*III site and extending increasing distances towards ARS1 were constructed (Fig. 2). These are called Δ 108– Δ 197; the numbers refer to the number of base pairs deleted. Restriction fragments of

plasmids containing these deletions were generated by cleaving at sites located 150–200 bp outside the deletion regions in such a way that the putative bend is approximately in the centre of the fragment. The mobility and *k* values of these DNA fragments are shown in Fig. 2. Both the wild-type (wt) and Δ 108 DNA fragments migrate more slowly than expected (31 and 32% more slowly, respectively). Δ 149 (which differs from Δ 108 by 41 bp) migrates normally. Thus at least some of the DNA that confers abnormal mobility lies between 108–149 bp on the map in Fig. 1.

Bent DNA tends to migrate more 'normally' at higher temperatures^{19,22}. ARS1 DNA fragments from the deletion plasmids were analysed as above except that the electrophoresis was performed at 38 °C instead of 25 °C. As shown in Fig. 2, the wild-type and Δ 108 fragments migrated faster at the higher temperature (lower *k* values). Relative mobilities of fragments lacking the 108–149 region were not affected by temperature. Thus, the effect of temperature on abnormal migration is consistent with the presence of bent DNA. Abnormal migration cannot be due to simple denaturation of DNA sequences, which would have the opposite effect with temperature. Inspection of the DNA sequence in the 108–149 region reveals the presence of six A_{3–5} stretches phased approximately every 10.5 bp (Fig. 3). This feature is characteristic of the bent DNA observed in kinetoplast DNA and λ replication origin^{6,8,21}. It has been shown²² that bending occurs at the 3' junctions of A tracts, and is maximized when the 3' A junctions are phased approximately every 10 bp.

Interestingly, DNA footprint analysis with partially purified yeast extracts demonstrates that a protein factor (ABFI) specifically binds to this region at residues 139–155 (Fig. 3). So far, five different binding sites for this factor have been identified in yeast DNA near ARS elements²⁴. The binding site at ARS1 contains an (A)₅ stretch directly in its centre and other binding sites contain (A)_{3–4} stretches. The bent conformation of the DNA at ARS1 may affect the affinity with which this factor binds DNA, similarly to SV40 T antigen where the nucleotides contacted by T antigen flank a DNA bend that is critical for binding⁷.

To determine whether the bent DNA region is important for ARS1 function *in vivo*, the effect of the five deletions shown in Fig. 2 on ARS activity was assayed. Each ARS fragment was

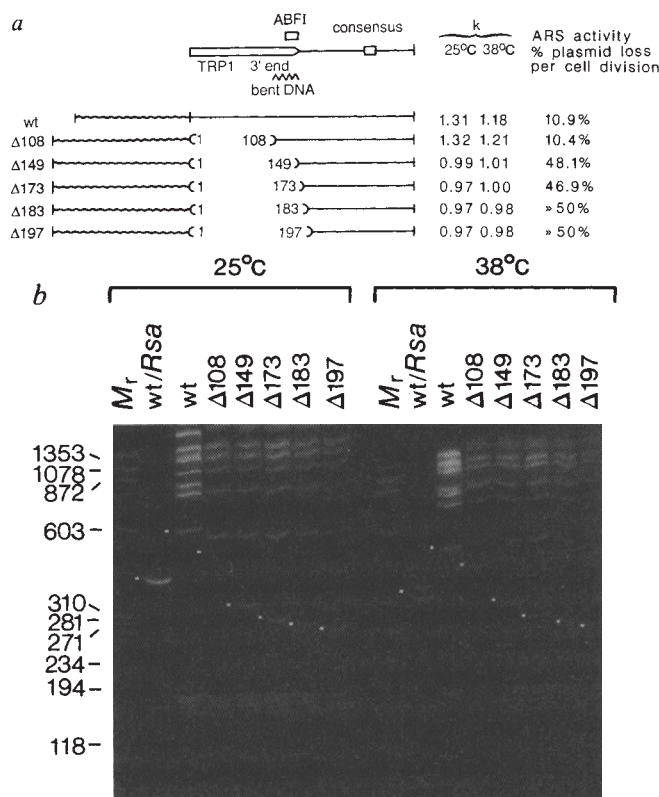


Fig. 2 Deletion analysis through the bent DNA region. *a*, The features of ARS1 and deletion fragments; *b*, mobility of digested deletion plasmids. *a*, All fragments were cloned into YRP14/CEN4²⁵. The wild-type (wt) fragment was inserted between the *Hind*III and *Eco*RI sites. Deletions begin at the *Hind*III site and extend to the nucleotides indicated. *Bam*HI octanucleotide linkers were added and the *Bam*HI/*Eco*RI fragments cloned into YRP14/CEN4. For bent DNA analysis wt DNA was cleaved with *Rsa*I and *Eco*RV which leaves 156 bp of pBR322 sequences (wavy line) adjacent to the *Hind*III/*Rsa*I ARS1 fragment. For the deletion plasmids, DNA was digested with *Rsa*I and *Sph*I which leave 187 bp of pBR322 sequences adjacent to ARS1 fragments. The lengths of the resulting fragments of interest are: wt 467 bp; Δ108, 396 bp; Δ149, 355 bp; Δ173, 331 bp; Δ183, 321 bp; Δ197, 307 bp. Electrophoresis was as in Fig. 1, at 25°C. DNAs were also digested with enzymes that remove the pBR322 sequences and analysed on the same gel. The wt fragment still migrates abnormally, but all deletion fragments Δ108–Δ197 now migrate with *k* values <1.05 (not shown). ARS activity for each plasmid was assayed by transforming the plasmids into yeast strain YNN318 (a/α *ura*3-52/*ura*3-52 *lys*2-801/*lys*2-801 *ade*2-101/*ade*2-101 *trp*1-Δ901/*trp*1-Δ901 *his*3-Δ200/*his*3-Δ200). Cells harbouring plasmids were transferred from selective yeast minimal plates containing glucose to yeast minimal media supplemented with galactose (Calbiochem) and grown for 16 h at 30°C (selecting for the plasmid) and an additional 10 h without selection, then plated on nonselective galactose plates containing limiting adenine; cells containing the plasmids appear pink while those lacking the plasmid are red²⁵. The frequency of $\frac{1}{2}$ sectored colonies among colonies containing a single plasmid ($n=296$ to 2,000) was measured to determine the rate of plasmid loss per cell division. Loss for wt and Δ108 plasmids was 4/1 1:0/2:0 segregation events while for Δ149 and Δ173 loss was >45/1 1:0/2:0 segregation events. Loss rates are very high for Δ183 and Δ197 constructs; all colonies appear homogeneously red. For Δ183 and Δ197 the number of plasmid-containing colonies plated was determined using selective glucose-containing plates. When similar ARS activity experiments were performed using glucose media, the relative stabilities of the wt and Δ108–Δ197 plasmids are unchanged; however, the overall plasmid stability is increased. This medium effect is consistent with our observations that ARS plasmids are more stable on richer carbon sources^{25,32}. *b*, Polyacrylamide gel showing mobility of the digested deletion plasmids. Left, DNAs analysed at 25°C for 16 h; right, identical gel electrophoresis at 38°C for 13 h. Lanes are as indicated and described above. Additional lanes are: *M*_r, ΦX174 *Hae*III digested DNA (sizes indicated in bp); wt/*Rsa*, *Rsa*I the *Hind*III/*Eco*RI ARS1 DNA digested with *Rsa*I. The 311-bp DNA fragment migrates coincident with a 356-bp DNA fragment at 25°C (the 356-bp DNA fragment migrated 11% more slowly than expected on this particular gel).

cloned into YRP14/CEN4, a centromere-containing plasmid lacking an efficient replicator, and the rate of plasmid loss was determined by a colony colour assay (Fig. 2)²⁵. Plasmids containing bent DNA (wild type and Δ108) are stable and lost at the

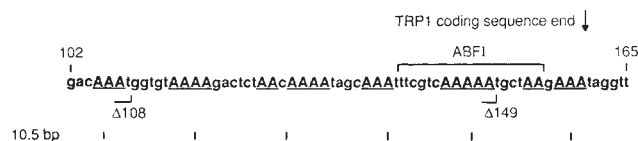


Fig. 3 DNA sequence in the bent DNA region of ARS1. All A residues that have neighbouring A residues are shown in bold type. 10.5-nucleotide intervals and the ABFI binding site are also indicated. A 25-bp region centred around this site is protected in DNase I protection experiments²⁰. The endpoints of deletions Δ108 and Δ149 are shown. The ARS1 consensus sequence is located at position 241–251.

rate of 10% per cell division. The Δ149 plasmid, which lacks most of the bent DNA and half of the ABFI binding site, is fivefold less stable (Fig. 2). Thus, the bent DNA region is important for ARS function *in vivo*. It remains to be determined whether this effect is due to the bent DNA sequence, to the binding site for ABFI, or to both.

Bent DNA has been previously discovered at the bacteriophage λ and SV40 replication origins^{6,7,26}. The discovery of such DNA at yeast ARS1 suggests that it may be important at chromosomal origins as well. Although the precise role of the bent DNA at ARS1 is presently unknown, the position of this region suggests several interesting possibilities. Since it is located at the 3' end of the *TRP1* gene, it may be important for transcription termination. Transcription through ARS1 has been found to reduce its activity (M.S. and R.W.D., unpublished observations). Thus one role of domains B and C might be to act as transcription termination segments to protect domain A. The bent DNA with its associated protein could serve an important role in this process. In addition, the region immediately adjacent to the ARS1 consensus element on a *TRP1* ARS1 plasmid (nucleotides 60–242) has been shown to be free of nucleosomes^{27,28}. Another function of the DNA bend might be to keep this region free of nucleosomes and readily accessible to the DNA replication apparatus and/or other proteins. The origin region of SV40 DNA has also been found to be free of nucleosomes^{29–31}.

We thank Ron Sapolsky for his help with sequence analysis and Roger Kornberg in whose laboratory some of these experiments were performed. This research was supported by NIH grant GM 21891.

Received 6 May; accepted 2 September 1986.

- Bossi, L. & Smith, D. M. *Cell* **39**, 643–652 (1984).
- Simpson, L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1585–1588 (1979).
- Ross, W., Schulman, M. & Landy, A. *J. molec. Biol.* **156**, 505–529 (1982).
- Stellwagen, N. C. *Biochemistry* **22**, 6186–6193 (1983).
- Chalberg, S. S. & Englund, P. T. *J. molec. Biol.* **138**, 447–472 (1980).
- Zahn, K. & Blattner, F. R. *Nature* **317**, 451–453 (1985).
- Ryder, K., Silver, S., DeLucia, A. L., Fanning, E. & Tegtmeyer, P. *Cell* **44**, 719–725 (1986).
- Marini, J. C., Levene, S. D., Crothers, D. M. & Englund, P. T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7664–7668 (1982).
- Hagerman, P. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4632–4636 (1984).
- Campbell, J. L. *Genet. Engng* **5**, 109–155 (1983).
- Newlon, C. S., Devenish, R. J., Suci, P. A. & Roffis, C. J. *ICN-UCLA Symp. molec. cell. Biol.* **22**, 501–516 (1981).
- Saffer, L. D. & Miller, O. L. *Molec. cell. Biol.* **6**, 1148–1157 (1986).
- Stinchcomb, D. C., Mann, C., Selker, E. & Davis, R. W. in *The Initiation of DNA Replication and Segregation of Yeast Chromosomes*. Vol. 3 (ed. Ray, D. S.) 473–488 (Academic, New York, 1981).
- Broach, J. R. *et al. Cold Spring Harb. Symp. Quant. Biol.* **47**, 1165–1173 (1982).
- Celniker, S. E., Sweder, K., Srien, F., Bailey, J. E. & Campbell, J. L. *Molec. cell. Biol.* **4**, 2455–2466 (1984).
- Kearsey, S. *Cell* **37**, 299–307 (1984).
- Koshland, D., Kent, J. C. & Hartwell, L. H. *Cell* **40**, 393–403 (1985).
- Srien, F., Bailey, J. E. & Campbell, J. L. *Molec. cell. Biol.* **5**, 1676–1684 (1985).
- Diekmann, S. & Wang, J. C. *J. Molec. Biol.* **186**, 1–11 (1985).
- Buchman, A., Kimmerly, W., Rine, J. & Kornberg, R., (manuscript in preparation).
- Wu, H. & Crothers, D. M. *Nature* **308**, 509–513 (1984).
- Koo, H. S., Wu, H.-M. & Crothers, D. M. *Nature* **320**, 501–506 (1986).
- Tschumper, G. & Carbon, J. *Gene* **10**, 157–166 (1980).
- Buchman, A. R. & Kornberg, R., (manuscript in preparation).
- Hietter, P., Mann, C., Snyder, M. & Davis, R. W. *Cell* **40**, 381–392 (1985).
- Zahn, K. & Blattner, F. R. *EMBO J.* **4**, 3605–3616 (1985).
- Thoma, F., Bergman, L. W. & Simpson, R. T. *J. molec. Biol.* **177**, 715–733 (1984).
- Thoma, F. & Simpson, R. T. *Nature* **315**, 250–252 (1985).
- Varshavsky, A. J., Sundon, O. & Bohm, M. *Cell* **16**, 453–466 (1979).
- Jakovits, E. B., Brutosomes, S. & Aloni, Y. *Nature* **285**, 263–265 (1980).
- Saragosti, S., Moyné, G. & Yaniv, M. *Cell* **20**, 65–73 (1980).
- Snyder, M. & Davis, R. W., (manuscript in preparation).

Flavonoid activation of nodulation genes in *Rhizobium* reversed by other compounds present in plants

J. L. Firmin, K. E. Wilson, L. Rossen
& A. W. B. Johnston

John Innes Institute, Colney Lane, Norwich NR4 7UH, UK

Bacteria of the genus *Rhizobium* specifically recognize and infect legume plants, resulting in a complex morphological and biochemical differentiation into nitrogen-fixing root nodules. In *R. leguminosarum*, whose host plant is the pea, and several other *Rhizobium* species, groups of linked plasmid-borne genes are involved in nitrogen fixation (*nif* genes) and nodulation (*nod* genes)^{1,2}. The *nod* genes are organized in several transcription units on the symbiotic plasmid pRL1J1 of *R. leguminosarum*; these comprise *nodABCIIJ*, *nodD* and *nodFE*³⁻⁶ with *nodD* being a positively acting regulatory gene whose product activates transcription of the other two operons which contain genes essential for the nodulation process^{5,7}. Although the precise biochemical functions of most of the pRL1J1 *nod* genes are unknown, it has recently been discovered that their expression requires not only the product of the regulatory *nodD* gene, but also a factor (or factors) present in pea root exudate^{5,7}, and similar results have been reported for *Rhizobium meliloti*⁸ and *Rhizobium trifolii*⁹. The finding of plant-produced products that 'signal' the *Rhizobium* offers the possibility of identifying one of the key biochemical reactions required for the infection process. Here we describe the characteristics of several plant-specified compounds that activate the transcription of the pRL1J1 *nodABCIIJ* and *nodFE* genes and present a novel phenomenon, namely that other, naturally occurring compounds can antagonize this induction.

For these studies the ability of different compounds to activate or to antagonize the activation of transcription of *nodABCIIJ* and *nodFE* was monitored by the use of gene fusions. Two *Rhizobium* test strains contained either the *nodC-lacZ* fusion plasmid pIJ1477 or the *nodF-lacZ* fusion plasmid pIJ1521, together with pIJ1518 which contains the cloned *nodD* gene^{5,7} (Fig. 1).

Preliminary studies showed the inducing compounds in pea root exudate to be soluble in water and methanol, dialysable and adsorbed by XAD-2 resin. High-voltage paper electrophoresis showed that the active component(s) behaved as a weak acid. Following high-voltage paper electrophoresis or paper chromatography (in several solvents) the region of biological activity co-migrated with material that absorbed ultraviolet light. These characteristics suggested that the active component(s) could be phenolic. HPLC fractionation of the exudate from germinating pea seeds showed four peaks of biological activity (Fig. 2) with one (peak A) being the most prominent. This was also the major peak in extracts from dried pea seeds and pea root exudate. Although exudate from seeds of alfalfa, a legume not nodulated by *R. leguminosarum*, could induce *nod* gene expression in this species, there were quantitative and

qualitative differences in the peaks of biological activity compared with those from peas (Fig. 2).

Electron impact mass spectroscopy (70 eV, 200 °C) of material in peak A from germinating pea seeds gave a complex pattern, with prominent ions at *m/e* 121 and 137. Although these might have been derived from simple phenolics, the presence of ions as large as *m/e* 256 suggested the presence of more complex phenolics. One such class of compounds are flavonoids; these are made by plants and some give fragments of *m/e* 121 and 137.

To confirm that flavonoid compounds were indeed inducers of *Rhizobium nod* genes, we took advantage of the fact that wild-type flowers of *Antirrhinum majus* contain many representatives of such molecules whereas a white flower colour mutant (termed *nivea*) which lacks chalcone synthase, contains none of these compounds¹⁰. A methanolic extract from red wild-type flowers induced both *nodC-lacZ* and *nodF-lacZ*, whereas an extract from the white *nivea* mutant flowers did not. Thus, one or more of the flavonoids in *A. majus* flowers induces *nod* gene expression in *R. leguminosarum*. To confirm this, several commercially available flavonoids known to be present in *A. majus* flowers, as well as other structurally similar flavonoids were tested as *nod* gene inducers (Table 1). Three flavones, three flavanones and one flavone glucoside induced *nod* gene transcription at low (10–100 nM) concentrations. However, flavones and flavanones lacking a hydroxyl group on the B-ring had no inducer activity; vitexin (8-*C*-glucosyl-apigenin) and the 7-*O*-rhamnoglucosides of hesperetin and naringenin also failed to act as inducers. Further, none of the isoflavones, flavanols or the flavone-3-glycoside tested had detectable inducer activity at concentrations as high as 10 µM.

Two of the active compounds, eriodictyol and apigenin-7-*O*-glucoside, had identical HPLC retention times to peak A obtained from peas. Evidence that peak A contained an *O*-glycoside of apigenin was that acid hydrolysis (2 M HCl, 100 °C, 2 h) of the material in peak A caused a reduction (by approximately 50%) of the inducing activity associated with this peak and generated a second, biologically active, ultraviolet-absorbing peak which had an identical retention time to authentic apigenin. (Further, authentic apigenin-7-*O*-glucoside was hydrolysed using the same conditions used for the hydrolysis of the material in peak A; following separation by HPLC, it was found that all the apigenin-7-*O*-glucoside was converted to apigenin.) Therefore peak A contained at least two biologically active compounds.

The individual compounds tested for activation of *nodF-lacZ* and *nodC-lacZ* were also examined for their ability to antagonize or augment induction by extracts from pea seed exudate. None of the compounds listed in Table 2 increased the levels of transcription of the *nodC-lacZ* or *nodF-lacZ* fusions. However, it was found that, at concentrations as low as 5 µM, the isoflavones daidzein and genistein and the flavanol kaempferol strongly inhibited the activation of the *nodC-lacZ* fusion (Table 2); similar results were obtained with these compounds using the *nodF-lacZ* fusion.

The second class of compounds tested as antagonists were analogues of acetophenone, which are structurally similar to the C-ring of the flavonoid *nod* gene inducers; several acetophenone analogues are known to induce transcription of the *vir*

Fig. 1 Location of pRL1J1 *nod* genes. Restriction sites used for the generation of the cloned *nodD* gene and the dimensions of recombinant plasmids used in this study. The locations and directions of transcription of the eight *nod* genes are shown. Restriction sites used for the generation of the cloned *nodD* gene in pIJ1518 (ref. 7) and of the *nodC-lacZ* and *nodF-lacZ* fusion plasmids pIJ1477 and pIJ1521 (refs 7 and 5 respectively) are indicated.

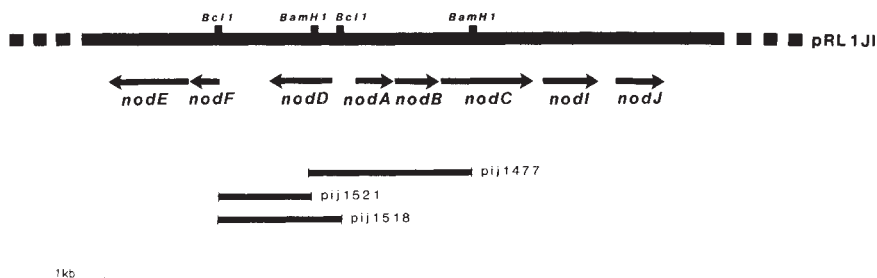


Table 1 Induction of *nodC-lacZ* expression in *R. leguminosarum* by selected flavonoids

Compound	Biological activity* at flavonoid concentrations of:			
	10 μ M	1 μ M	100 nM	10 nM
Flavones				
4',5,7-trihydroxyflavone (apigenin)	5,163	4,649	541	59
3',4',7-trihydroxyflavone	6,814	1,069	109	64
3',4',5,7-tetrahydroxy- flavone (luteolin)	5,120	1,437	77	73
Flavone glucoside				
4',5,7-trihydroxyflavone 7-O-glucoside (apigenin 7-I-glucoside)	5,964	3,642	239	55
Flavanones				
4',5,7-trihydroxyflavanone (naringenin)	4,552	3,154	100	58
3',5,7-trihydroxy-4'- methoxyflavanone (hesperitin)	9,207	6,409	1,036	250
3',4',5,7-tetrahydroxy- flavanone (eriodictyol)	7,144	5,066	523	75

The *lac* fusion strain KH1122 pIJ1518 pIJ1477⁷ was grown to a density corresponding to $A_{600} = 0.3$ prior to exposure to the flavonoid for 15 h at 29 °C with shaking. The bacteria were then assayed for β -galactosidase activity as described previously⁷. The following flavonoids had no detectable activity in the concentration range 10 μ M–10 nM:

Flavones: 5,7-dihydroxy-; 5,7-dihydroxy-; 3',4'-dimethoxy-; 7-hydroxy-4'-methoxy-; 3,5,7-trihydroxy-; 3,4,5,7-tetrahydroxy-.

Flavone glycosides: 4',5,7-trihydroxyflavone 8-C-glucoside; 3',4',5,7-tetrahydroxyflavone 3-O-rutinoside.

Flavanone: 5-hydroxyflavanone.

Isoflavones: 4',5,7-trihydroxy- (genistein); 4',7'-dihydroxy.

Flavanone glycosides: 3',5-dihydroxy-4'-methoxyflavanone 7-O-rhamnoglucoside; 4'-5-dihydroxyflavanone 7-O-rhamnoglucoside.

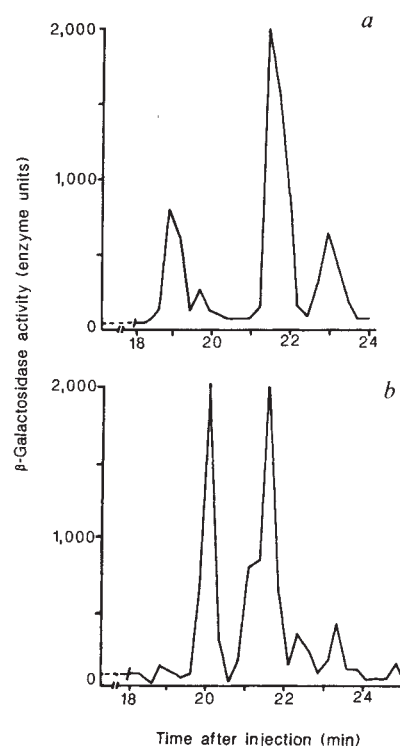
* Biological activity is expressed as units of β -galactosidase activity. In the absence of any added inducer the background level of expression was ~60 units.

genes of *Agrobacterium tumefaciens*¹¹. None of these compounds induced *nod* gene expression but acetovanillone and 4-hydroxyacetophenone (at 50 μ M) strongly inhibited induction of *nodC-lacZ* and *nodF-lacZ* by pea seed exudate (Table 2). At higher concentrations (1 mM) several other related phenolic compounds also strongly inhibited *nod* gene activation (Table 2). None of the compounds affected the growth rate of the cells.

The results presented here show clearly that certain naturally occurring flavonoids activate *nodD*-dependent transcription of the *R. leguminosarum nodFE* and *nodABCII* transcriptional units and that this induction is antagonized by other flavonoids or simpler phenolics present in plants.

In general, flavones and flavanones (Fig. 3), which are hydroxylated in both the A- and B-rings, acted as inducers. All flavonoids tested in which the C3 position was substituted with a hydroxyl group (flavonols), a sugar (flavone-3-glycoside) or phenyl group (isoflavonoids) were inactive. At least one B-ring hydroxyl group appears to be necessary for inducing ability (for example 4',5,7-trihydroxyflavone is active but 5,7-dihydroxyflavone is not). The B-ring hydroxyl can be at either (or both) the 3' or 4' position; thus hesperitin (3'-hydroxy-4'-methoxy), naringenin (4'-hydroxy) and eriodictyol (3',4'-dihydroxy) are all active.

The presence of a B-ring methoxy at the 4' position appears to enhance potency, since hesperitin, the 4'-OMe-ether of eriodictyol, is more active than eriodictyol itself. However, no biological activity was detected when a 3'-methoxy- or 3',4'-dimethoxy- was the only substituent on the B-ring (see Table 1). The glucosylated flavone apigenin-7-O-glucoside (but not apigenin-8-C-glucoside) acts as an inducer. It is possible, though, that the 7-O-glycoside is deglycosylated by *Rhizobium* to form apigenin,

**Fig. 2** HPLC analysis of *nod* gene inducers present in the seed exudate from pea (a) and alfalfa (b).

Methods. Seeds of pea or alfalfa were sterilized with aqueous sodium hypochlorite (50%) and were then soaked in sterile distilled water for 3 days with gentle shaking. The liquid was removed, and after centrifugation (5,000g, 15 min) the supernatant was concentrated *in vacuo* (45 °C) and was applied to a preparative reverse phase (C8) HPLC column (12 cm $\times$ 0.7 cm). The column was eluted at 2 ml per min with a linear gradient of methanol-water, from 10% to 99% methanol at 3.2% per min, and fractions of 5 ml were collected. After evaporation, a portion of each fraction was assayed for *nod* gene-inducing ability using *nodC-lacZ* fusion⁷. Active fractions were combined, concentrated and applied to an analytical HPLC column (Brownlee Aquapore RP300, 25 cm $\times$ 0.46 cm fitted with a 3 cm $\times$ 0.46 cm guard cartridge). The column was eluted at 1 ml per min with a linear gradient of methanol-water (0–3.3 min, 10% methanol, isocratic; 3.3–24.7 min, 10–85% methanol, linear gradient; 24.7–28.1 min, 85% methanol, isocratic; 28.1–32.0 min, 85–100% methanol, linear gradient). Fractions (0.25 ml) were tested for the ability to induce *nod* gene expression as described above. Similar results were obtained with the *nodF-lacZ* fusion instead of *nodC-lacZ*.

itself an inducer of *nod* gene expression. Moreover, the distinction between active and inactive flavonoids might reflect their differential uptake or metabolism rather than their (presumed) interaction with the regulatory *nodD* gene product.

Activators of *R. leguminosarum nod* gene transcription were found not only in pea, which is nodulated by this species of *Rhizobium*, but also in extracts of alfalfa, clover and cowpea plants which are not nodulated by *R. leguminosarum*. Thus host range specificity cannot be determined solely on the basis of *nod* gene inducers such as the active flavonoids described here.

Further, the flavone luteolin was shown to be the major active component present in alfalfa seeds for the induction of *nodABC* transcription in *R. meliloti*¹². This flavone also activated *R. leguminosarum nod* genes (Table 1). HPLC analysis of extracts of peas and alfalfa seed showed that they both contained several molecules with inducing activity. In peas the most prominent peak on HPLC (peak A in Fig. 2) contains at least two compounds (this would explain the complexity of the mass spectrum). Taken together the evidence from chromatography, mass spectroscopy and acid hydrolysis strongly suggested that the material in peak

Table 2 Effect of selected flavonoids and acetophenone analogues on *nod* gene expression in *R. leguminosarum*

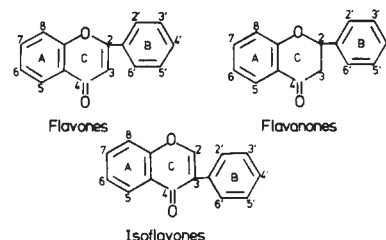
Compound	% Inhibition at flavonoid concentrations of:					
	1,000 μ M	500 μ M	100 μ M	50 μ M	10 μ M	5 μ M
5,7-dihydroxy,3',4'-dimethoxyflavone	ND	0	0	0	0	0
5,7-dihydroxyflavone	ND	24	20	3	0	0
4',5-dihydroxyflavanone	ND	0	0	0	0	0
7-O-rhamnoglucoiside (naringin)	ND	0	0	0	0	0
4',5,7-trihydroxyflavone 8-C-glucoside (vitexin)	ND	0	0	0	0	0
3,3',4',5-tetrahydroxy-7-methoxyflavone (rhamnetin)	ND	83	68	30	0	0
3,4',5,7-tetrahydroxyflavone (kaempferol)	ND	82	82	65	69	60
4',7-dihydroxyisoflavone (daidzein)	ND	85	88	89	85	75
4',5,7-trihydroxyisoflavone (genistein)	ND	91	82	58	78	81
4-hydroxy-3,5-dimethoxybenzoic acid (syngic acid)	83	39	31	29	23	ND
4-hydroxy-3,5-dimethoxy-benzaldehyde (syngaldehyde)	88	37	22	23	23	ND
4-hydroxy-3,5-dimethoxycinnamic acid (sinapinic acid)	89	60	28	15	22	ND
4-hydroxy-3-methoxyacetophenone (acetovanillone)	95	95	91	85	20	ND
4-hydroxy-3,5-dimethoxyacetophenone (acetosyringone)	96	94	95	46	5	ND
3,5 dimethoxyacetophenone	92	69	14	21	19	ND
4-hydroxyaceto-phenone	94	95	94	91	41	ND

The *lac* fusion strain KH1122 pIJ1518 pIJ1477⁷ was pregrown to an absorbance (at 600 nm) of 0.15 prior to exposure to the compound under test, in the presence of a fixed amount of natural inducer (obtained from germinating pea seeds) equivalent to 2,000 units of β -galactosidase activity. The bacteria were shaken for 15 h at 29°C and were then assayed for β -galactosidase activity as described previously⁷. ND, not determined.

A contained apigenin-7-O-glucoside plus another flavonoid which had the same chromatographic behaviour as eriodictyol. Further, Redmond *et al.*¹³ found that in clover the major activator of *R. trifolii nod* gene transcription was 7,4'-dihydroxyflavone.

The finding that other naturally occurring compounds made by plants inhibit *nod* gene induction by flavonoids is novel. Those flavonoids which were most potent in antagonizing induction had hydroxylation patterns in the A- and B-rings similar to those of the inducing flavonoids described here but different from the inducing molecules in having a substituted C3 position, as in the isoflavonoids genistein and daidzein and the flavonol kaempferol. Flavones lacking a B-ring hydroxyl (for example, 5,7-dihydroxyflavone) were much less inhibitory. Given the structural similarities between antagonists and the inducing molecules it is possible that they act competitively either at the level of transport or within the bacterial cell. The substitution pattern of the acetophenone analogues tested also had a marked effect on their potency as inhibitors of *nod* gene induction; the most potent of this group were those which had a substitution pattern [(3-methoxy, 4-hydroxy-) or 4-hydroxy-] similar to that found in the B-ring of the flavonoids that induced *nod* gene transcription.

There is an apparent paradox in the fact that isoflavonoids were among the most potent antagonists of *nod* gene transcription; *a priori*, it might have been thought that isoflavonoids,

**Fig. 3** Skeleton and numbering nomenclature for the flavonoids mentioned in this study.

which are found predominantly in the family Leguminosae¹⁴, would be good candidates for compounds that activate the expression of *Rhizobium nod* genes.

The concentrations required for antagonism of induction by, for example, the isoflavonoid daidzein were one to two orders of magnitude greater than the concentration required for *nod* gene induction by the most potent of the activating molecules. However, isoflavonoids are present in germinating legumes at concentrations as high as 20 μ M even in healthy plants¹⁵. It seems likely therefore that, during the normal nodulation process, the regulation of *Rhizobium nod* genes is subject both to activation and to the antagonism of activation by different plant-specified molecules. Indeed, compounds were found in pea root exudate which antagonize the induction of the *nodC-lacZ* fusion by the inducing compounds in peak A. This was demonstrated by fractionating root exudate by the HPLC and adding individual fractions of exudate to a strain carrying a *nodC-lacZ* fusion which had been grown in the presence of the inducer (C. A. Shearman, personal communication). The host plant legume can regulate the number of nodules it contains on its root system. This has been clearly demonstrated by the fact that mutants of soybean¹⁶ and of pea¹⁷ have been isolated which contain many times more nodules than wild-type. Further, when a recombinant plasmid containing the cloned pRL1JI *nodABC* genes, transcribed constitutively from a vector promoter, was introduced into a wild-type strain of *R. leguminosarum* nodulation of peas was severely inhibited¹⁸ indicating that continuous high-level transcription of these genes is deleterious for the infection process. It will be of interest to see if the control of nodule number depends upon the extent to which the *nod* genes are expressed, this being determined by the relative concentrations of the inducers and anti-inducers which are present in the rhizosphere and in the plant root tissue.

We thank Keith Parsley for mass spectrometry, Enrico Coen for providing *Antirrhinum* plants and David A. Hopwood for useful discussions about the manuscript.

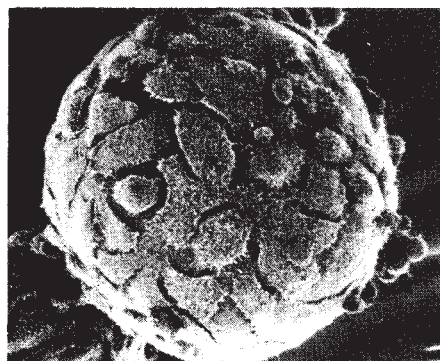
Received 23 May; accepted 14 July 1986.

- Long, S. R. in *Plant Microbe Interactions*, Vol. 1, *Molecular and Genetic Aspects* (eds Kosage, T. & Nester, E. W.) (Macmillan, New York, 1984).
- Rossen, L. *et al.* in *Oxford Survey of Plant Molecular and Cell Biology* (ed. Miflin, B. J.) (in the press).
- Downie, J. A., Knight, C. D., Johnston, A. W. B. & Rossen, L. *Molec. gen. Genet.* **198**, 255-262 (1985).
- Rossen, L., Johnston, A. W. B. & Downie, J. A. *Nucleic Acids Res.* **12**, 9497-9508 (1984).
- Shearman, C. A., Rossen, L., Johnston, A. W. B. & Downie, J. A. *EMBO J.* **5**, 647-652 (1986).
- Evans, I. J. & Downie, J. A. *Gene* (in the press).
- Rossen, L., Shearman, C. A., Johnston, A. W. B. & Downie, J. A. *EMBO J.* **4**, 3369-3373 (1985).
- Mulligan, J. T. & Long, S. R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6609-6613 (1985).
- Innes, R. W. *et al.* *Molec. gen. Genet.* **201**, 426-432 (1985).
- Spiribille, R. & Forkmann, G. *Phytochemistry* **21**, 2231-2234 (1982).
- Stachell, E. S., Messens, E., van Montague, M. & Zambryski, P. *Nature* **318**, 624-629 (1985).
- Peters, G. M. & Long, S. R. *J. Bact.* (in the press).
- Redmond, J. W. *et al.* *Nature* (in the press).
- Harborne, J. B. in *Chemotaxonomy of the Leguminosae* (eds Harborne, J. B., Boulter, D. & Turner, B. L.) 31-71 (Academic, London, 1971).
- Partridge, J. E. & Keen, N. T. *Phytopathology* **66**, 426-429 (1976).
- Carrol, B. J., McNeil, D. L. & Gresshoff, P. M. *Pl. Physiol.* **78**, 34-40 (1985).
- Jacobsen, E. & Feenstra, W. J. *Pl. Sci. Lett.* **33**, 337-344 (1984).
- Knight, C. D., Rossen, L., Robertson, J. G., Wells, B. & Downie, J. A. *J. Bact.* **166**, 552-558 (1986).

Notes for the laboratory

From cell culture products to an expert program, this week highlights innovative and useful ideas.

Ventrex Laboratories, Inc. has recently introduced **neutrally buoyant glass micro-carrier beads** (Reader Service No. 100). Typical Ventreglas beads have a specific gravity of only 1.03 making the vigorous stirring speeds that retard cell growth, but are required with heavier beads, unnecessary. Ventrex's beads are polystyrene coated with glass, and are therefore



The Ventreglas bead: unsinkable

capable of withstanding cleaning by trypsinization, chromic acid, or alcohol. Ventreglas beads are also autoclavable. The beads come in a variety of diameters ranging from 90–210 μm , and a 25 g sample is available for \$100 (US).

Statistical Programs' new expert computer system, **EXPERTIMENTAL DESIGN**, eases the selection of **statistical designs for data analysis** (Reader Service No. 101). Through a question-and-answer format, the program ranks the suitability of 13 designs, ranging from factorial to Box-Behnken analysis, to the user's experiment. Uncertainty in the user's answers is dealt with by assigning desirabilities and manipulating them according to fuzzy set theory. Accompanied by a user's manual including a tutorial, the programmed floppy diskette sells for \$295 (US), and is written for 256K IBM PC/XT/AT.

Mycoplasma contamination need not plague your cell cultures and influence your experimental results, now that the American Type Culture Collection is offering their new **mycoplasma detection service** (Reader Service No. 102). ATCC's procedure costs \$50 (US) each for 1–10 samples, and utilizes both a direct

and an indirect assay to determine if suspect cultures are affected. The direct assay involves a total incubation time of 28 days, according to ATCC, while the indirect assay (using a fluorochrome staining procedure) can be accomplished within a week.

The membrane-based **culture plate inserts** for epithelial cell culture available from Millipore allow *in vitro* cultures to closely mimic cells grown *in vivo* by providing cells with both serosal and mucosal contact with media (Reader Service No. 103). Millicell-HA inserts come presterilized in blister packages for convenience, and are produced in a 12 mm configuration to fit 24-well plates, as well as a 30 mm configuration appropriate for 6-well

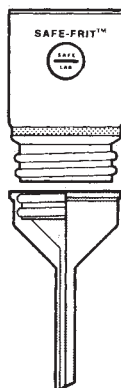
plates. Available since June 1986 in the United States, Millicell-HA inserts are now available in the United Kingdom, and are priced at £50–70 (UK) (depending on size) for a box of 50.

Taconic, Inc. eliminates worries over **clinically silent infections in laboratory animals** with its free bimonthly *Lab Animal Health Report* (Reader Service No. 104). Each edition will feature a different strain, and will outline the battery of tests to which Taconic mice and rats are subjected, along with ratios of incidence of infection. Available to any researcher using mice and rats, Taconic's reports will allow experimenters to exclude subclinical infection from other influential variables in the interpretation of results. □

ADVERTISEMENTS

SAFE-FRIT™ FUNNELS

Safe-Lab's NEW fritted glass funnels feature a removable, unbreakable polypropylene stem just below the fritted glass disc. This unique design eliminates the danger of broken glass funnel tips and permits easy cleaning from both sides. Removable stem allows you to set the funnel top on counter or directly inside an oven or desiccator for drying, without the danger of tipping or spilling. Polypropylene has excellent chemical resistance and threads will not leak, even at high-vacuum.



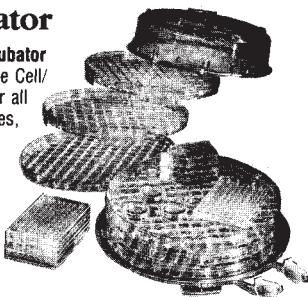
Call or write for a **FREE** catalog:
Safe-Lab, Inc., P.O. Box 1290,
Santee, CA 92071
(619)448-9100

Reader Service No.28

Versatile Tissue Culture Incubator

Modular Incubator Chamber (the Cell/Sphere™) for all tissue cultures, aerobic or anaerobic.

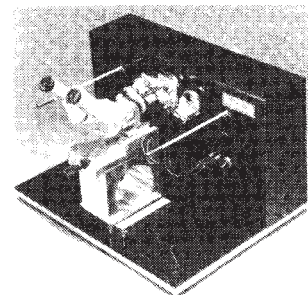
Low Cost



Eliminates evaporation loss and contamination. Easy to use – flush with gas mixture, seal, place at appropriate temperature. Stackable, see-through. Quantity discounts. **billups-rothenberg, inc.**, p.o.b 977, Del Mar, CA USA 92014. Call 619-755-3309 or contact your Flow Laboratories representative.

Reader Service No.7

MICRO-TOOL MAKING! then you need the **SINGER MICROFORGE**

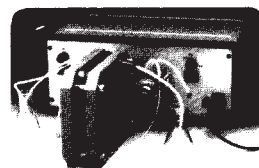


Write to: Singer Instrument Co Ltd, Treborough Lodge, Roadwater, Watchet, Somerset TA23 0QL, England. Tel: (0984) 40226. Telex. 337492 Comcab G.

Reader Service No.50

LIPOSOMES

REPRODUCIBLE – HOMOGENEOUS –
EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5–10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.35

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

The changing graduate labour market

Richard Pearson

Competition between employers for the best UK graduates is becoming intense.

As the new academic year gets under way, final-year undergraduates have to face up to the twin pressures of final exams and job hunting. While 1987 promises to be another buoyant year for graduate recruitment, the traditional pattern of the annual 'milkround' is coming under increasing strain as employers seek the 'best' graduates at a time when the universities are producing fewer graduates.

Thus, alongside generally booming job prospects, the graduate labour market is showing clear evidence of fragmentation, reflecting a more general trend in the labour market as the market becomes 'freer' and employers compete intensively for the 'best' and for specialist skills, offering increasing salary differentials and special premiums for quality staff, while those with less marketable skills suffer at the bottom end of the labour market chasing low paid and often temporary jobs, or experiencing bouts of unemployment. As a result, while 1985 saw the largest ever number of university graduates entering employment, unemployment six months after graduation remained at significant levels, 8.6 per cent, although this was somewhat lower than in previous years (Table 1) and compares favourably with unemployment in the population at large.

This year well over 100,000 students graduated with first degrees, with the uni-

otherwise not available for work six months after graduation. The 33,500 university graduates going into employment were split in roughly equal proportions between public service/education, industry, and finance/commerce, with the latter having shown the most growth during the past five years.

Interestingly, while the numbers entering jobs in industry had grown steadily between 1980 and 1984, they fell slightly in 1985. For this year the forecasts have been of a five per cent increase in overall vacancies with particular growth being shown in the fields of transport, government, public utilities, financial and other services. The decline in engineering and other manufacturing vacancies is expected to continue.

Traditionally the vast majority of graduate vacancies have been advertised and filled through the annual 'milkround', when employers visit the campuses in the spring term, announcing their vacancies and carrying out first interviews. Nowadays, however, an increasing number of employers are starting their visits before Christmas and making job offers much sooner in order to be able to select the best staff. The numbers doing this have risen from 12 to 17 per cent of recruiters in the last three years (*The Cost Effectiveness of the Graduate Milk Round*, IMS, 1986). At the same time more and more vacancies are being advertised through the summer fairs, with over one in three recruiters using them in 1986, some to compensate for shortfalls from the milkround but also because some new vacancies appear in the summer months. More students also appear to be delaying job-searching to concentrate on their exams, secure in the knowledge that good jobs will be available in the summer months. Thus while the milkround remains the dominant recruitment method, the graduate labour market is becoming less concentrated in time.

Another sign of the fragmentation of this part of the labour market is the increasing range of starting salaries being offered. In the early 1970s virtually all graduates received similar starting salaries with only small distinctions being made between different types of employers and little distinction being drawn between graduates from different disciplines. By the late 1970s, premiums were being paid to engineers in shortage disciplines, and with the demise of incomes policies, more emphasis on market forces and increased competition, the diversity of starting

salaries has increased markedly. For example, in 1986 the median graduate starting salary was about £8,000 per year, but with an upper decile of £9,000 and lower decile of £6,750. The annual rate of increase at the top of the range of 9 per cent was double that at the bottom, further widening the gap.

Last week's 'Big Bang' in the City of London had also had a highly publicised impact on graduate salaries, with 1986 seeing many examples of new city recruiters offering starting salaries of between £10,000 and £13,000 per year — in several

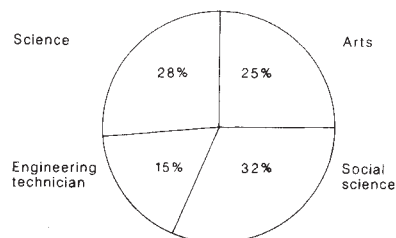


Fig. 1 Subject distribution, all institutions. (Source: University Grants Committee.)

cases even more. Looking to 1987 it seems clear that many more financial institutions will be offering starting salaries in between £12,000 and £14,000 a year, which will be well ahead of the likely average of about £9,000 for most jobs. With golden hellos also being paid (one merchant bank, for example, giving 'bursaries' of £2,000 to let graduate recruits 'travel in the summer vacation'), decisions about jobs and careers can become even more complex.

The key for job hunters is to look beyond the short-term rewards of salary and also think about job interest, training and experience, and the longer-term career development prospects, which may be gained through several firms. While these latter factors can often act as compensation for a lower starting salary, the disturbing feature of many jobs on offer is that the lower salaries are often combined with the poorer training and career prospects, further emphasizing the fragmentation of the labour market into a smaller elite of highly rewarded and well trained staff, pulling away from the broader range of jobs available to the majority, and unemployment or underemployment facing a significant minority. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Falmer, Brighton BN1 9RF, UK.

Table 1 Graduate unemployment

	Unemployment six months after graduation (%)	Entering permanent employment
1979-80	9.0	29,882
1980-81	11.3	28,523
1981-82	13.5	29,602
1982-83	11.6	32,506
1983-84	9.6	33,942
1984-85	8.6	34,717

Source: University Grants Committee

versities accounting for two out of three graduates, and the polytechnics and colleges one in three. The majority graduated in arts and social science subjects, with under 15 per cent being engineers and technologists, while a further 28 per cent were scientists (Fig. 1). Not all of those graduating, however, seek to, or actually enter the labour market, about one in four going on to further academic work or formal training (although only 3,200 went into teacher training in 1985, a fall of over 40 per cent since 1981) while about one in six were unemployed or